Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations	1. Thesis and Dissertation Collection, all items

1964

Technical notes on probability and statistics for business decisions

Magee, William; Wood, Robert D.

Monterey, California: U.S. Naval Postgraduate School

http://hdl.handle.net/10945/13249

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

http://www.nps.edu/library

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

NPS ARCHIVE
1964
MAGEE, W.

TECHNICAL NOTES ON PROBABILITY AND STATISTICS FOR BUSINESS DECISIONS

WILLIAM C. MAGEE
ROBERT D. WOOD

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5101

Library
Naval Postgraduate School
Monterey, California

TECHNICAL NOTES ON PROBABILITY AND STATISTICS FOR BUSINESS DECISIONS

by

William C. Magee

Commander, United States Navy

and

Robert D. Wood

Commander, United States Navy

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE

IN

MANAGEMENT

United States Naval Postgraduate School

Monterey, California

1964

NPSARCHIVE
1964
MAGEE, W.

Library
U. S. Naval Postgraduate School
Monterey, California

TECHNICAL NOTES ON PROBABILITY AND STATISTICS FOR BUSINESS DECISIONS

by

William C. Magee

and

Robert D. Wood

This work is accepted as fulfilling
the thesis requirements for the degree of
MASTER OF SCIENCE
IN
MANAGEMENT
from the

United States Naval Postgraduate School

ABSTRACT

Schlaifer's text, though considered one of the best for students of Business Statistics, nevertheless is difficult for many students to understand. This is particularly true of those who are returning to college studies after a considerable lay-off. The authors of this Research Paper felt that future students might be aided by technical notes in which the material is presented in a somewhat down-to-earth manner that would be easier to understand. The Technical Notes supplement Schlaifer's text and cover the same concepts; that is, decision under uncertainty; the use of probabilities based directly on experience; binomial sampling, the normal distribution and sampling of measured values.

INTRODUCTION

The purpose of these Technical Notes is to supplement the basic theory and explanations of the basic text, Introduction to Statistics for Business Decisions, Robert Schlaifer, McGraw-Hill Book Company, 1961, York, Pennsylvania.

These Technical Notes are primarily for

(1) The average student who may have never been exposed to statistical theory.

(2) The student who has been away from school for a number of years.

(3) The student who finds Schlaifer a bit too rapid for good digestion.

The titles to the chapters are congruent with the titles given in the basic text. An attempt has been made to introduce problems to further illustrate the basic theory presented by Professor Schlaifer in order that the student may gain greater facility in problem setups, and gain a deeper sense of familiarity and feeling for statistics as it applies to the everyday application both in the Navy and outside of it.

Professors Freund and Williams describe the relationship of business statistics in modern business:

> . . . most elementary textbooks stressed the collection and sources of data and the presentation of numerical data in tables, charts, pictures, and maps.

> Although the importance of statistics in this phase of business cannot be denied, there has been a pronounced change in attitude toward the subject: business statistics is now viewed as providing quantitative bases for arriving at well-informed decisions with respect to all matters connected with the operation of a business. In short, it is now recognized that, in its broadest sense, business statistics includes also the methods and inferences needed to provide an adequate flow of quality raw materials, to select and evaluate the performance of both machines and personnel, to design products and maintain their quality, and to evaluate new methods of production, advertising, and selling the goods and services of industry, both in the immediate present and in the near and distant future.[1]

There are several types of statistics, such as, descriptive statistics, inductive statistics, vital statistics, and quantum statistics. This Technical Notes are primarily concerned with inductive statistics. Descriptive statistics is understood to mean the treatment of any numerical data which does not involve generalizations. When predictions, estimations, or generalizations are made, we are speaking of inductive statistics. For example, take five light bulbs of brand A and five light bulbs of brand B and burn them continuously until they burn out. We can easily determine the average lifetime of the five bulbs of each brand. Suppose brand A produced the following re-

---

[1]Professor John E. Freund, and Professor Frank J. Williams. Modern Business Statistics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1958. p. 2.

TABLE OF CONTENTS

CHAPTERS

1. The meaning of Probability
2. Expected Value and Utility
3. Random variables and Probability Distributions
4. Opportunity loss and the Cost of Uncertainty
5. Incremental Analysis
6. Measures of Location; Fractiles and Expectations, Linear Profits and Costs
7. Assessment of Probabilities by Smoothing Historical Frequencies
8. Conditional Joint, and Marginal Probability
9. The Binomial Distribution
10. Statistical Decision Rules and their Error Characteristics
11. Evaluation of Statistical Decision Rules in Terms of Expected Loss
12. Revision of Probabilities in the light of New Information
13. Suspension of Judgment and Summarization of Information
14. Measures of Variability. Variance and Standard Deviation
15. The Normal Distribution and the Normal Approximation to the Binomial Distribution

16. The Central Limit Theorem and Large Sample Theory

17. Statistical Decision Rules with Normal Sampling

18. Expected Loss with Normal Sampling

INTRODUCTION

The purpose of this syllabus is to supplement the basic theory and explanations of the basic text, Introduction to Statistics for Business Decisions, Robert Schlaifer, McGraw-Hill Book Company, 1961, York, Pennsylvania.

This syllabus is primarily for

(1) The average student who may have never been exposed to statistical theory.

(2) The student who has been away from school for a number of years.

(3) The student who finds Schlaifer a bit too rapid for good digestion.

The titles to the chapters in this syllabus are congruent with the titles given in the basic text. An attempt has been made to introduce problems to further illustrate the basic theory presented by Professor Schlaifer in order that the student may gain greater facility in problem setups, and gain a deeper sense of familiarity and feeling for statistics as it applies to the everyday application both in the Navy and outside of it.

Professors Freund and Williams describe the relationship of business statistics in modern business:

> . . . most elementary textbooks stressed the collection and sources of data and the presentation of numerical data in tables, charts, pictures, and maps.

> Although the importance of statistics in this phase of business connot be denied, there has been a pronounced change in attitude toward the subject: business statistics is now viewed as providing quantitative bases for arriving at well-informed decisions with respect to all matters connected with the operation of a business. In short, it is now recognized that, in its broadest sense, business statistics includes also the methods and inferences needed to provide an adequate flow of quality raw materials, to select and evaluate the performance of both machines and personnel, to design products and maintain their quality, and to evaluate new methods of production, advertising, and selling the goods and services of industry, both in the immediate present and in the near and distant future.[1]

There are several types of statistics, such as, descriptive statistics, inductive statistics, vital statistics, and quantum statistics. This syllabus is primarily concerned with inductive statistics. Descriptive statistics is understood to mean the treatment of any numerical data which does not involve generalizations. When predictions, estimations, or generalizations are made, we are speaking of inductive statistics. For example, take five light bulbs of brand A and five light bulbs of brand B and burn them continuously until they burn out. We can easily determine the average lifetime of the five bulbs of each brand. Suppose brand A produced the following re-

---

[1] Professor John E. Freund, and Professor Frank J. Williams, Modern Business Statistics. (Englewood Cliffs: Prentice-Hall, Inc., 1958) p. 2.

sults:

$$\frac{985 + 863 + 1024 + 972 + 746}{5} = 918 \text{ average hours/bulb}$$

and Brand B:

$$\frac{892 + 1071 + 993 + 929 + 785}{5} = 934 \text{ average hours/bulb.}$$

The computations thus far are in the realm of descriptive statistics. If it is concluded from this data that brand B is superior to brand A, then a generalization has been made. This is characteristic of inductive statistics. It is readily apparent from a cursory inspection of the data, that the generalization might or might not hold true. If we were dealing with 6 light bulbs, the data from the sixth light bulb of each brand could reverse the averages, i.e., the average lifetime of brand A is now greater than the average lifetime of brand B. In either case, risk is involved of making the wrong generalization. The analysis, evaluation and control of chances associated with this type of risk is one of the main functions of inductive statistics.[2]

---

[2] Ibid. pp. 3-4

CHAPTER I

1. THE MEANING OF PROBABILITY

1.1 Description

Not long ago, the owner-manager of an automobile dealership said to me, "My business has prospered because of luck. Every move I have made has been a gamble, some have been won and some have been lost but if one risks nothing, he'll not have a business very long." This businessman is a dealer in automobiles. Further discussion with him concerning his decisions for ordering cars by color, model, interiors, parts, etc., revealed that almost all decisions along these lines were made intuitively either by himself or his sales manager. He said that he is continually faced with the problems like: How much advertising is the optimum amount? (Four hundred dollars for a two minute plug on a TV station is approximately equal to the profit in one car sale) How many skits should he buy? At what time of day are his potential customers most likely to be watching TV, etc.? How much gas should be put into the cars sold to customers for good will? These are a few examples in this business where some kind of a probability model might be an excellent aid to the solution of some of these problems that are solved intuitively now.

Everyone faces the future with uncertainty in private,

public or business life, in every profession, and in every government of the world. For example, we are constantly aware of the uncertainties that can occur during a naval operation. How many destroyers are needed to successfully complete a mission? A destroyer cannot steam without feedwater; what are the chances of her evaporators breaking down? How great a risk is acceptable of not completing a mission? How reliable are the Polaris Missiles and how can we estimate their reliability? How good is our estimate? What are the risks as determined by tests and what risks are acceptable? These are a few of the problem areas that can be examined in the light of statistical theory. There are literally thousands of applications of it. To introduce elementary probability, the die model offers an excellent framework for thinking about it.

1.2 The Die Model[1]

Can you think of a die with other than six sides? One with two sides may be constructed by the intersection of two parallel planes with a sphere. A three-sided die may be constructed by the intersection of a sphere and three skew planes, etc. With some imagination one can expand this thought process visualizing a four-sided, a

[1] Paul G. Hoel, Elementary Statistics (John Wiley & Sons, Inc., 1960), pp. 36-37.

five-sided, the normal six-sided die, a ten-sided, twenty-sided, or "infinitely-sided" die (sphere). What probabilities would you assign to each face of each kind of die coming up in a roll? You would most likely assign to the two-sided die, $\frac{1}{2}$, to the three-sided 1/3, to the four-sided 1/4, etc., and to the infinitely-sided die $1/\infty$ or zero. This concept is easy to see if you visualize rolling a giant sphere on a large flat surface. What probability would you assign to its coming to rest on any particular point on its surface? Assuming that it would rest on one molecule, there are almost an infinite number of molecules on which it could rest. The probability that it would rest on any particular one would be almost $1/\infty$ or zero.

Take a die with six faces, intuitively a probability of 1/6 would be assigned to any number 1 to 6 coming up on the roll. This assessment could be incorrect if for example, the die had a five on every face, or it was loaded, filed on one side or corner, or shaved. A dishonest die would be assigned different values to the probability of each face coming up on a roll.

Assuming that all the possible experimental outcomes can be expected to occur with the same frequency, then the probability of occurence of each event can be defined as follows:

The probability that an event A will occur is the ratio of the number of outcomes that correspond to the occurrence of A to the total number of possible outcomes.[2]

This definition fits the case of the honest die. As an illustration, take a pair of honest dice. What possible combinations are there? The possible outcomes are listed in Table 1.

TABLE 1

11	12	13	14	15	(16)
21	22	23	24	(25)	26
31	32	33	(34)	35	36
41	42	(43)	44	45	46
51	(52)	53	54	55	56
(61)	62	63	64	65	66

Each digit in Table 1 is the number on a die. The circled values indicate the combinations that total 7. In accordance with the definition, a probability of 6/36 would be assigned to event A = 7 coming up on a pair of dice. There are 36 possible outcomes and 6 outcomes corresponding to A. This may be written P(A) = 6/36. Let B = 11. The probability of B is 2/36. Let C = 11 or 7. Then P(C) = 8/36.

---

[2] Ibid. pp. 36-37

Suppose the dice are dishonest and all the 1's are replaced by 2's. The shape of the dice has not been changed so there would be no change to the probabilities assigned to each face coming up on a roll. This is a case of the more general type, one in which the various outcomes of an experiment are not expected to occur with the same relative frequency. In the more general case, the definition is modified:

The probability that an event A will occur is defined as the sum of the expected relative frequencies for the possible outcomes that correspond to the occurrence of A. For example, let A = 7. The student should verify

$$P(A) = 2/36 + 1/36 + 1/36 + 2/36 = 1/6$$

because these numbers are the expected relative frequencies corresponding to the four favorable outcomes, i.e., 25, 34, 43, and 52. This result is the same as P(A) in the first illustration. For the combinations 16 and 61, has been substituted the combinations 25 and 25. Let B = 4. The only possible outcome now is 22 since there are no 1's on the dice. Then

$$P(B) = 4/36$$

because there are four ways of making B with these dishonest dice. The result is not the same as that for the previous experiment of rolling two honest dice.

> This procedure of assigning expected relative frequencies to each of the possible outcomes of an experiment, whether they are assumed to be equal or not, is part of the process of constructing a theoretical model for the actual experiment. If the assigned relative frequencies are close to the relative frequencies that will be obtained when the experiment is repeated a large number of times, then the probability of an event A calculated on the basis of those assigned relative frequencies will be close to the relative frequency with which the event A will occur when the experiment is repeated a large number of times. Thus, in rolling two dice, the model that one usually chooses is the one in which all 36 outcomes are assigned the same relative frequency. Since experience with dice has shown that this is a realistic model, it will usually be found that the observed relative frequency of getting a total of, say, 7 points will be close to the probability value of 1/6 if a large number of rolls is made. It is the statistician's privilege to assign any consistent relative frequencies he desires to the possible outcomes of his experiment; however, he will naturally attempt to make his model as realistic as possible so that his resulting probabilities of events will be realistic and therefore useful for solving practical problems.[3]

It should also be noted that on the basis of historical data obtained in the recent past, the statistician will modify his original probability assignments to make his model more realistic. The more realistic the probability model, the more likely the conclusions drawn from it about the future will be realistic.

1.3 The Addition Rule.[4]

If A and B are mutually exclusive events, this means if A occurs, then B can not occur. For example, if in

---

[3] Ibid. p. 38.
[4] Ibid. p. 40.

the dice illustration a 7 comes up in a roll, 11 can not come up in the same roll. The number of outcomes corresponding to the occurrence of A + B is equal to the sum of the number of outcomes corresponding to the occurrence of A and the number of outcomes corresponding to the occurrence of B. Then

$$P(A + B) = P(A) + P(B).$$

For more than two mutually exclusive events the formula may be extended, e.g.,

$$P(A + B + C) = P(A) + P(B) + P(C)$$

As an illustration with two dice again let A, B, and C, be the events of getting a total of exactly 10 points, 11 points, and 12 points, respectively. From Table 1, the probabilities of those events occurring are given by $P(A) = 3/36$, $P(B) = 2/36$, and $P(C) = 1/36$. The probability that at least one of these mutually exclusive events will occur is given by

$$P(A + B + C) = 3/36 + 2/36 + 1/36 = 1/6.$$

1.4 The Sum of the Probabilities of all events is 1.00

Take the case of a two-sided die or coin. The probability of getting H(heads) in the first toss is $\frac{1}{2}$. The probability of getting T(tails) the first toss is $\frac{1}{2}$. $\frac{1}{2} + \frac{1}{2} = 1.00$. One can also look at this probability concept from the point of view of one event. The P(H) i.e., (probability of getting heads) on the first toss is

$\frac{1}{2}$, and the probability of $\bar{H}$ (not heads) not getting heads is 1.00 - $\frac{1}{2}$ = $\frac{1}{2}$. Extend this line of thinking to the six-sided die. The P(5) coming up on a throw is 1/6, and the P ($\bar{5}$) (probability of <u>not</u> 5) is 1.00 - 1/6 = 5/6.

1.5 <u>Problems</u>

1. Find the P of drawing from a contract bridge deck of cards:

   (a) An ace

   (b) Four aces without replacement

   (c) A spade

   (d) Not a spade

   (e) Not an ace

   (f) Two aces without replacement on two draws

   (g) A red card or a jack on one draw

Answers. (a) 4/52 (b) 4/52 x 3/51 x 2/50 x 1/49

(c) 1/4 (d) 3/4 (e) 12/13

(f) 4/52 x 3/51 (g) 28/52

CHAPTER 2

Expected Value and Utility

2.1 Illustrative Definitions

2.1.1 Conditional value-The value of an act on condition that some event occurs, eg: the value to a businessman of a contract on condition that he gets the contract; or the value to a businessman (cost to him) of submitting a contract proposal on condition that he does not get the contract; also the value to a businessman of a contract on condition that he does nothing. For example, a contract with a $35,000 gross profit which cost $10,000 to prepare for would have a $25,000 value on condition of being won, a minus $10,000 value on condition of being not won and a zero value on condition of not submitting a proposal at all.

2.1.2 Expected value--A weighted average of all the conditional values; the weights applied to the conditional values are the probabilities that the alternative events will occur, eg: the probability that the businessman will get the contract and the probability that he will not get the contract. In simplest terms it is the old fashioned arithmetic average.

2.2 The expected monetary value-of an act can be found by

CHAPTER 2

Expected Value and Utility

2.1 Illustrative Definitions

2.1.1 Conditional value-The value of an act on condition that some event occurs, eg: the value to a businessman of a contract on condition that he gets the contract; or the value to a businessman (cost to him) of submitting a contract proposal on condition that he does not get the contract; also the value to a businessman of a contract on condition that he does nothing. For example, a contract with a $35,000 gross profit which cost $10,000 to prepare for would have a $25,000 value on condition of being won, a minus $10,000 value on condition of being not won and a zero value on condition of not submitting a proposal at all.

2.1.2 Expected value--A weighted average of all the conditional values; the weights applied to the conditional values are the probabilities that the alternative events will occur, eg: the probability that the businessman will get the contract and the probability that he will not get the contract. In simplest terms it is the old fashioned arithmetic average.

2.2 The expected monetary value-of an act can be found by

constructing a payoff table as follows:

Table 2.1 Expected monetary Value of Making the Proposal

EVENT	Probability	Conditional Value	Expected Value
Get contract	.5	$25,000	$12,500
Do not get contract	.5	- 10,000	- 5,000
	1.0		$ 7,500

The Act of submitting a proposal for this $35,000 gross profit contract has an expected monetary value of $7,500.

Frequently expected value may be negative. For example, what is the expected value to each player when they have an even $10 bet on the toss of a coin? One player has a .5 probability of winning $10 and the other has a .5 probability of losing $10. In this case the expected monetary value will be zero: $10(.5) - $10(.5) = 0. Suppose that for some foolish reason player A feels sure he will win and is willing to bet $20 against player B's $10; what is the expected monetary value of the toss to player A? $10(.5) - $20(.5) = - $5. To player B? -$10 (.5) + $20(.5) = $5.

2.3 In addition to problems of a monetary nature the concept of expected value is also applicable in other areas such as inventory and sales forecasting. A military problem of inventory control of an electron tube for which there might be 12,000 gross requirements would have an expected

value of 5,800 units as follows:

Table 2.2 Computation of an Expected Inventory Requirement

EVENT	Probability	Conditional Requirement	Expected Requirement
Hot war	.25	10,000	2,500
Cold war	.60	5,000	3,000
Normal peacetime	.15	2,000	300
	1.00		5,800

Conditional requirements in this case are based on the anticipated number of equipments in use over a period of time for which there is a requirement for this particular tube, and the probabilities are estimates assigned to the occurrences of hot war, cold war or normal peacetime, respectively.

2.4 The concept of expected monetary value provides a means by which we can consider the various consequences of two or more events. Some of these consequences may be positive and some may be negative, and in summation they tend to balance each other somewhat. Expected monetary value may be likened to a man standing with his right foot on a block of ice and the left one on a warm stove. The experience is neither all hot nor all cold, but one or the other will likely predominate. Looking at it alternatively, if he had just come in from the cold winter weather, he might shift his weight to his left foot; but in a hot summer situation he might shift to the right.

Similarly, if in our previous example Player A wants to raise the game to $200 against player B's $100, player B might not be willing to continue the game any longer even though the expected monetary value is quite favorable to him--he cannot afford to lose $100. But Player A, being a well-heeled speculative individual, is willing to risk his money for a much smaller gain. This indicates that not only do individuals have different utilities for money, but each person's individual utility varies at different levels (just as the man had differing utilities for the ice and stove under varying circumstances).

When a businessman is faced with a decision which has one of two possible consequences such as the aforementioned contract problem, expected monetary value may or may not be an effective guide to action. If his financial condition is such that he can afford to take a $10,000 loss if he does not get the contract, then he might be quite willing to accept that risk in order to gain the opportunity of a $25,000 net profit. On the other hand, if his financial condition were marginal, he could ill afford the consequence of losing $10,000 if he did not get the contract and the expected monetary value of the contract would be zero to him for he would not submit a proposal at all.

Therefore in making the decision on whether or not to use expected monetary value as a guide to action, the

businessman should reduce his problem to a decision between (1) an act which is certain to result in receipt of a certain amount of cash and (2) an act which will result in either the best or the worst of all consequences of the decision problem. In the case illustrated above this would be a choice between (1) $0 and (2) either a profit of $25,000 or a loss of $10,000 for an expected monetary value of $7,500. If the stakes are so large that the businessman does not want to make the decision in this manner, then expected monetary value would not be a valid guide for his action and another decision-making guide must be found. This is done by the use of individual utility curves.

2.5 Development of Utility Curves

Utility is defined as the power to satisfy human wants. The objective of the individual is held to be maximization of the total utility he can achieve with his limited resources of time, money, and effort. The rationality of the individual is defined in terms of the utilization he makes of his scarce resources to achieve this end of maximization of utility.[1] It is derived from the economic theory of diminishing marginal utility. The more you have of a commodity, including money, the less desireable is the possession of one more unit of it. To own a Cadillic is most desireable and a second one

[1]David W. Miller and Martin K. Starr, Executive Decisions and Operations Research, (Englewood Cliffs, N.J.: Prentice Hall Inc., 1961) p. 20.

might be quite nice but by the time you got to the third one, its subjective value (utility) to you is quite a bit less than the second and substantially less than the first. The same is true of steaks, ice cream, money, etc.

A utility curve is a graphic indication of how an individual will react when faced with certain types of decisions of a monetary nature. The basic text gives a detailed and thorough explanation of how these curves are developed. The example in the text is reviewed here because it provides a concise explanation of utility.

Utility curves have some limitations:

1) A utility curve applies only to <u>one</u> individual in a <u>particular</u> type of situation and not to all types of situations with which this individual might be faced;
2) utilities cannot be added-the utility of one consequence cannot be added to that of another to obtain the utility of both together;
3) utility curves are not guides to social action-because one man cannot afford to lose as much as another as indicated by a more conservative utility curve is no justification to tax the man who can afford to be more risky.

The factors considered in deriving an individual's utility curve reveal that any particular curve is a subjective

indication of that individual's behavior pattern when he is faced with certain types of decisions-his attitude toward risk, profit and loss.

Figure 2-1 shows the utility curves of four individuals who are faced with decisions that can result in a series of consequences from losing $10,000 to making a profit of $10,000:

Individual A--is probably a marginal operator who must have a very good probability of making a profit before he is willing to risk any loss at all in order to gain a profit. He is so conservative that he is willing to pay a substantial amount for insurance against a loss even though the probability of loss is small.

Individual B--is probably either a speculator or an individual with substantial resources who is willing to risk much in order to gain an even bigger profit. Even though the probability of making a large profit is small, he would take a chance on it in preference to a small profit on a sure thing.

Individual C--is somewhat of a moderate type and is between A and B in his attitude toward risk. He uses expected monetary value as a guide

to action.

Individual D--takes a very conservative attitude toward risk when the probability of profit is small but as the probability of success increases, he becomes more and more willing to take a chance.

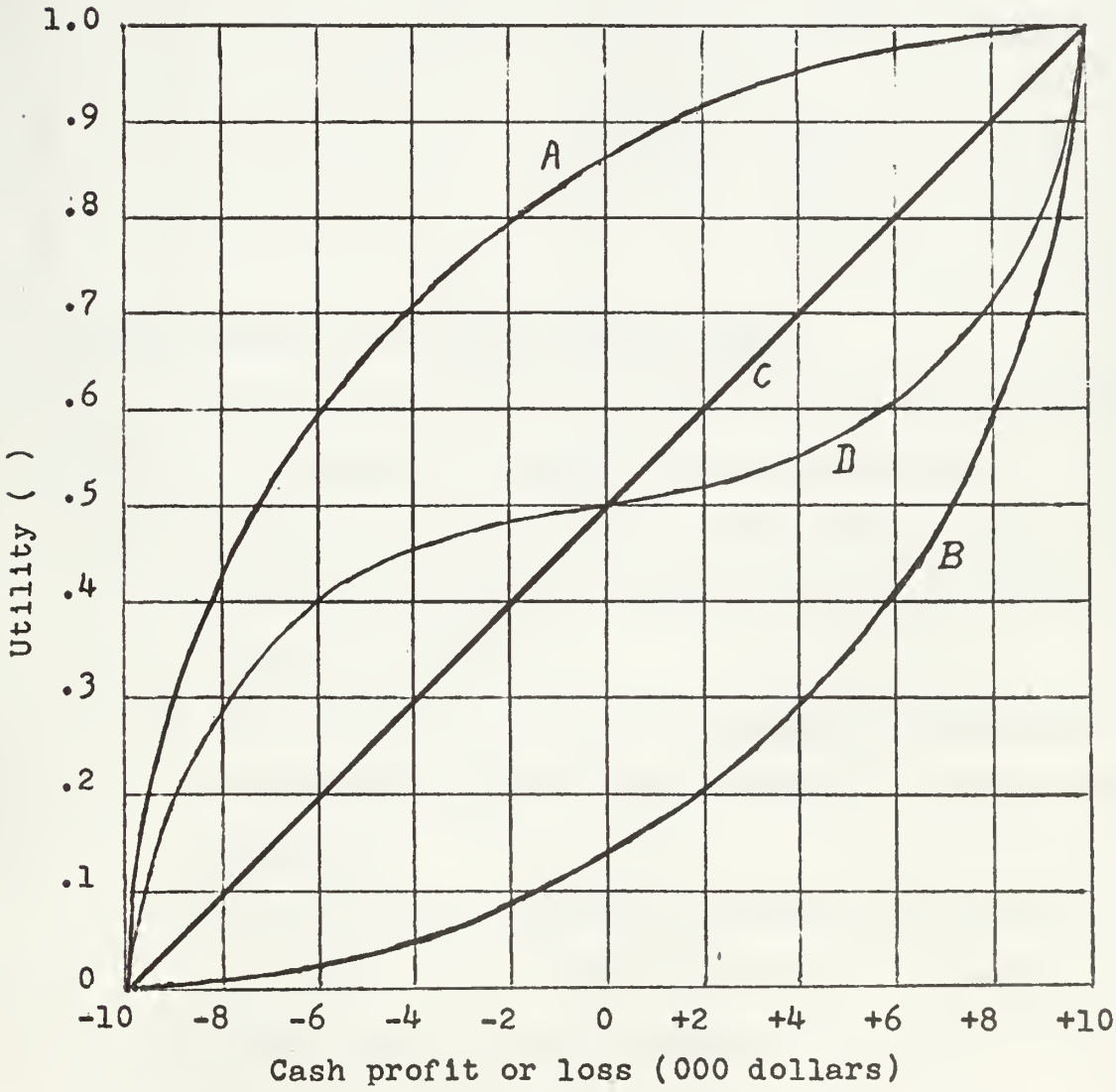


Figure 2.1. Utility of money for four different individuals

Table 2.3. Cash equivalent of a $10,000 profit or loss contract for four different individuals.

Probability π of $10,000 profit	Cash equivalent for Individual A	B	C	D
.85	0	9,700	7,000	9,500
.75	-3,000	9,000	5,000	8,500
.50	-7,000	7,000	0	0
.25	-9,000	3,000	-5,000	-8,500
.15	-9,700	0	-7,000	-9,500

Utility curves have the following uses:

1) They can be used to compute the cash equivalent of a particular contract for a businessman.
2) They indicate how much a businessman might be willing to pay for insurance against certain types of losses.
3) They indicate the point at which a businessman is indifferent about either accepting or rejecting a contract.
4) They provide a method of analysis for determining whether or not a businessman should accept or reject certain types of contracts that are offered him.
5) They indicate whether or not expected monetary

value can be used as an effective guide to action. (In the case of Businessman C, it can.)

Problems on expected Monetary Value

1. A new furniture store in a growing community is considering purchase of a certain type of wall-to-wall carpeting for resale. Based on the number of new homes being built, they have assigned the following probabilities to the market for wall-to-wall carpeting for the next year.

Sales in Sq. Ft.	Probability
6,000	.1
10,000	.2
20,000	.5
45,000	.2

If the store purchases 5,000 sq. yds. at $3.00/yd., what must the selling price be in order to realize a profit of $3,000?

Solution:

Sales in Sq Ft	Probability	Expected Sales
6,000	.1	600
10,000	.2	2,000
20,000	.5	10,000
45,000	.2	9,000
	1.0	21,600

Gross income: $15,000 + $3,000 = $18,000

Dividing gross income by expected number of sales,

$$\frac{18,000}{21,600} = \$ .83/\text{ sq. ft. or } \$7.50/\text{sq. yd.}$$

2. A naval officer has just inherited $10,000. He is a speculative type who is considering "investing" his windfall in an afternoon at the races. But he is also interested in a more conservative long range investment. After much careful study of the Racing Form, he estimates that he has about a 50-50 chance of winning $8,000 or losing $4,000 on an afternoon of betting. He has another opportunity to make a $10,000 investment in real estate. He estimates, that if business conditions are good, that within a year the land will be developed and he can make a $5,000 profit, but if no development takes place, he will lose $5,000 on a subsequent resale when he is transferred. He estimates a .75 probability that the land will be in demand for development and a .25 probability of no development. If he loses at the races, he cannot make the real estate investment because he has no other source of funds. If he uses expected monetary value as a guide to action, what choice should he make? What is the expected monetary value of each act?

Solution:

There are three possible acts each with an expected monetary value as follows:

Act I-An afternoon at the races:

Event	Probability	Expected Monetary Value
Win	.5	.5 x $8,000 = $4,000
Lose	.5	.5 x -4,000 = -2,000
	1.0	$2,000

Act II-Forego the races and invest in real estate:

Event	Probability	Expected Monetary Value
Area is developed	.75	.75 x $5,000 = $3,750
Area is not developed	.25	.25 x -5,000 = -1,250
	1.00	$2,500

Act III-An afternoon at the races and invest in real estate:

Event	Probability	Expected Nonetary Value
Lose $4,000	.5	.5 x -$4,000 = -$4,000
Win $8,000/ lose $5,000	.5 x .25 = .125	.125 x 3,000 = 750
Win $8,000/ gain $5,000	.5 x .75 = .375	.375 x 13,000 = 4,875
	1.000	$1,625

Since Act II has the highest expected monetary value, he should forego the races and invest in real estate.

3. An individual is considering investing $10,000 in a ship-building company. If there is peace and increased international trade the investment will return 5% but if there is a war it will return only 4% due to excess profits taxes. If the investor assigns 40% probability to peace and 60% to war, what is the conditional value of the investment under each condition? Should he make an alternative investment that offers a sure thing at 4⅕%?

Solution:

EVENT	Probability	Conditional Value	Expected Value
Peace	.4	$10,000 x .05 = $500	$500 x .4 = $200
War	.6	10,000 x .04 = 400	400 x .6 = 240
			$440

A sure thing at 4⅕% would result in an expected monetary value of $450. Since this is better than $440, it would be the better investment for an individual using expected monetary value as his decision criterion.

4. An investor owns a fleet of trucks which he has leased to a manager-operator for a fee based on a share of the profits. If successful, the present value of future rentals is estimated at $100,000, but if not successful, the present value would be $20,000. The manager-operator has offered to buy

the trucks outright from the investor for $68,000. If the investor uses expected monetary value as a guide to action what probability should be assigned to success in order for him to be indifferent about selling? If he considers the probability between success and failure about even, should he sell?

Solution:

Event	Conditional Value	Probability	Expected Value
Success	$100,000	P	$100,000 P
Failure	20,000	1-P	20,000 (1-P)
		1	$ 80,000 P + $20,000

The expected monetary value of the sale is $68,000. If the investor is to be indifferent between selling and renting, the two actions must have the same expected monetary value,

$$\$80,000\ P + \$20,000 = \$68,000$$

$$P = .6$$

If he considers the probability of success .5, and to be indifferent requires a probability of success of .6, then the probability of success is not high enough and he should sell. Suppose the offer to purchase outright had been $59,000, should he sell? No. Why not?

Problems on Utility

5. Assume that the below data is derived from the Risky Investment Company.

Utility Π	Equivalent Consequence
0	- $50,000
.3	- 40,000
.4	- 30,000
.45	- 20,000
.5	0
.55	20,000
.6	30,000
.7	40,000
.1	50,000

a) Draw a utility curve and compute the expected utility for the following investments. Indicate which, if any, the Risky Company should make:

Investment X			Investment Y		
Event	Probability	Consequence	Event	Probability	Consequence
P	.30	-$40,000	T	.10	+$48,000
Q	.35	+ 45,000	U	.50	+ 10,000
R	.15	0	V	.40	- 35,000
S	.20	- 12,000			

b) How much should the Risky Company be willing to pay for insurance against a $25,000 loss if the probability of loss occurring is .25?

c) For what probability of a profit would the Risky Company be indifferent about accepting an investment on which they could either make or lose $20,000?

d) For what probability of a profit would they be indifferent about accepting an investment on which they could either make $30,000 or lose $15,000?

Answers:

a) The expected utility of investment X is .539, investment Y is .482. They should therefore accept investment X but not Y.

b) $500

c) .5

d) .23

Note: Answers to the above problems may differ somewhat due to differences in plotting and graph interpretation.

CHAPTER 3

Random variables and Frequency Distributions

3.1 Histograms and right and left tail distributions

An important concept to get well fixed in your mind at this point is that of the histogram and the right and left tail distributions of the random variable. Upon this basic foundation will be built many interesting and important ideas. Let's examine for a moment what a random variable is. We shall define it and then show some examples and variations of its form and use. The symbol $\tilde{z}$ (read z-tilde) is the specific value that the variable may take. $P(\tilde{z} = z)$ is the short way of saying, "the probability that a certain elementary event obtains for which the random variable $\tilde{z}$ takes the value z."[1] In the following example, the random variable is the number of heads i.e., $\tilde{H}$, and it can take any value H, meaning $\emptyset$ heads, 1 head, 2 heads, or 3 heads.

To illustrate the use of these symbols, a theoretical frequency distribution is generated and shown below. Take a fair coin and toss it in sequences of 3. There are eight possible outcomes:

HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
3	2	2	2	1	1	1	0

[1] Harry V. Roberts, Statistics for Management, (University of Chicago, 1961) pp. 2-8.

Notice that below each combination is described the number of heads in each series. By thus defining the number of heads, we have implicitly defined the number of tails.

Since each toss of a coin is an independent event, it can be tossed three times or three coins tossed once and derive the same data. The P of the eight combinations are:

$$P(H = 3) = 1(\tfrac{1}{2}x\tfrac{1}{2}x\tfrac{1}{2}) = 1/8$$

$$P(H = 2) = 3(\tfrac{1}{2}x\tfrac{1}{2}x\tfrac{1}{2}) = 3/8$$

$$P(H = 1) = 3(\tfrac{1}{2}x\tfrac{1}{2}x\tfrac{1}{2}) = 3/8$$

$$P(H = 0) = 1(\tfrac{1}{2}x\tfrac{1}{2}x\tfrac{1}{2}) = 1/8$$

$$\text{Total P} = 1.0$$

This is a discrete distribution meaning that it is impossible to get a fraction number of heads i.e., $1\frac{1}{2}$ H. So, the results may be displayed as follows:

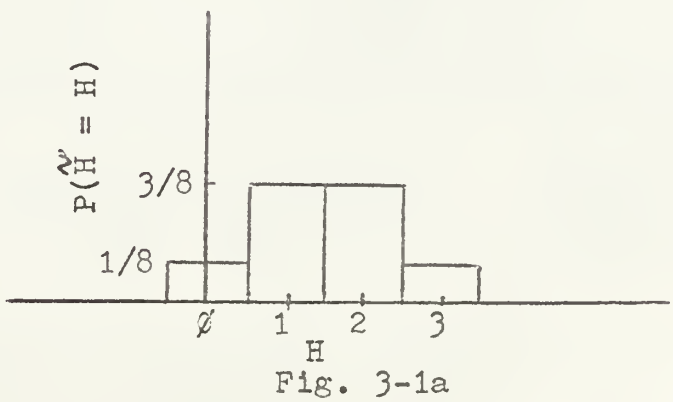
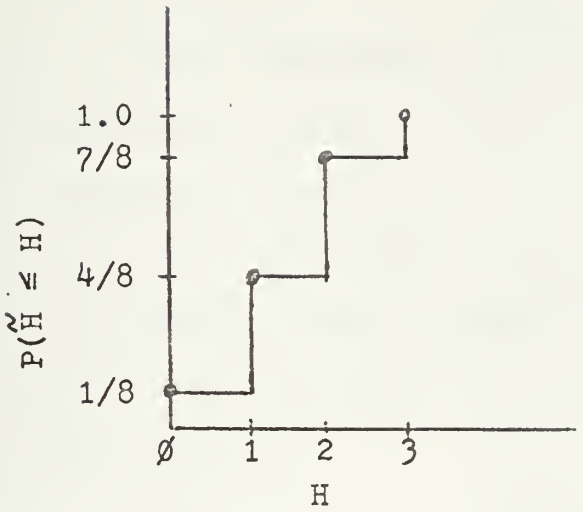


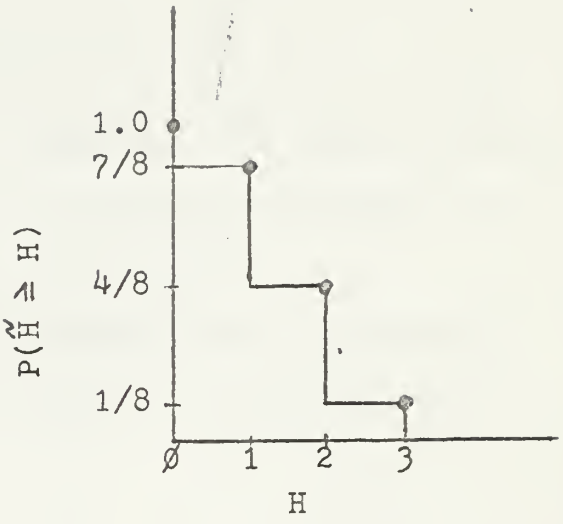
Fig. 3-1a

Note: P(H = H) Heads is the random variable and can be either Ø, 1, 2, or 3.



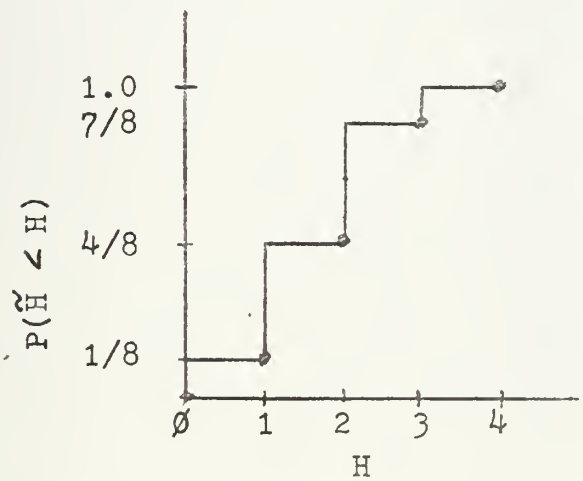
Left-tail distribution

Fig. 3-1b



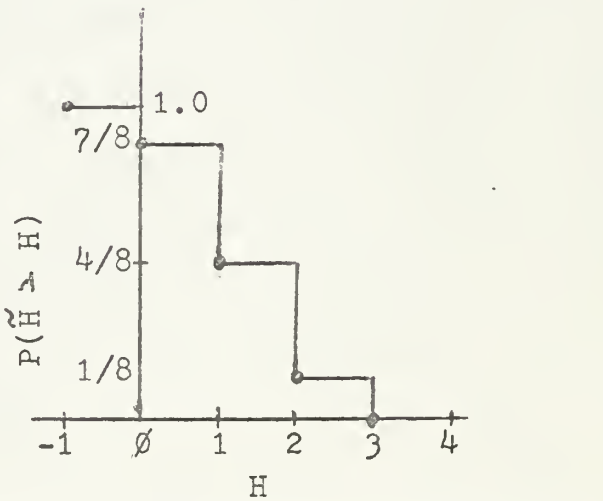
Right-tail distribution

Fig. 3-1c



Left-tail distribution

Fig. 3-1d



Right-tail distribution

Fig. 3-1e

$P(\tilde{H} \leq 3) = 1.0$ means the probability of getting either three heads, two heads, one head, or zero heads in three tosses of a coin is 1.00.

$P(\tilde{H} \geq \emptyset) = 1.0$ means the probability of getting $\emptyset$ heads, one head, two heads, or three heads in three tosses of a coin is 1.00.

From the derivation of P(H = 0, 1, 2, or 3) notice the coefficients 1, 3, 3, 1; these are the coefficients of the third series of the binomial expansion of the form

$$(a + b)^3 = \underline{1}a^3 + \underline{3}a^2b + \underline{3}ab^2 + \underline{1}b^3$$

For any series these coefficients may be developed by Pascal's triangle the first few of which are given in fig. 3-4.

Fig. 3-4

Pascal's Triangle

				1			$(a + b)^0$	
			1		1		$(a + b)^1$	
		1		2		1	$(a + b)^2$	
	1		3		3		1	$(a + b)^3$
1		4		6		4		1	$(a + b)^4$
1	5		10		10		5	1	$(a + b)^5$
1	6	15		20		15	6	1	$(a + b)^6$
1	7	21		35		35	21	7	1	$(a + b)^7$

Note: Each number in the table is the sum of the two numbers above on its left and right.

Now consider the two-sided die (coin) again and take $(a + b)^3$ expanded; notice the congruency. From paragraph 3.2 we have:

$$1\,HHH + 3HHT + 3HTT + 1TTT$$

$$H^3 + 3H^2T + 3HT^2 + T^3$$

$$a^3 + 3a^2b + 3ab^2 + b^3$$

The implicit probabilities are:

$$P(a) = \tfrac{1}{2}$$

$$P(b) = \tfrac{1}{2}$$

We may now step into the realm of the ten-sided die and consider it binomially. Take one particular face called A. What is the probability of that face coming up?

$P(A) = 1/10$. The $P(\bar{A})$ --the probability of <u>not</u> <u>A</u> coming up--is 9/10.

We now have an analogous situation compared with the two-sided die except for the implicit probabilities associated with a ten-sided die. Symbols adopted now for the sake of clarity are A which represents one side of the ten-sided die and N which represents <u>not</u> <u>A</u>. Now write this in binomial form and expand say, for five tosses. What are the probabilities of each event?

(1) $(A + N)^5$

(2) The coefficients from Pascal's triangle

1 5 10 10 5 1

(3) Then

$$A^5 + 5A^4N^1 + 10A^3N^2 + 10A^2N^3 + 5A^1N^4 + N^5$$

(4) The probabilities for each event follows:

$$(1/10)^5 + 5(1/10)^4(9/10)^1 + 10(1/10)^3(9/10)^2 + 10(1/10)^2(9/10)^3 + 5(1/10)^1(9/10)^4 + (9/10)^5$$

(5) Perform the indicated operation and:

1×10^{-5} = P(5A's)

45×10^{-5} = P(4A's)

810×10^{-5} = P(3A's)

7290×10^{-5} = P(2A's)

32805×10^{-5} = P(1A's)

59049×10^{-5} = P(ØA's)

1.00000 = Total P as it should be.

We have just developed a model that can be used throughout the remainder of Schlaifer's textbook and will be referred to occasionally as we proceed through the remainder of this syllabus.

To recap briefly, the concept of a multisided die from one with two sides to one of an infinite number of sides has been introduced. The relationship of probability to number of sides has been shown. The binomial expansion has been related to the possible outcomes of a series of tosses and the mathematical probabilities inherent therein.

Problems

1. Develop and plot the right tail cumulative binomial distribution for a series of five tosses of a fifty-sided fair die.
2. An urn contains 5 black and 10 white balls. What is (a) the P(B)? (b) The P(B) on two successive draws with replacement after each draw? (c) What is P(2B) without a replacement? (d) The P(B) followed by white with replacement? (e) The P(W) followed by black without replacement?

Answers.

2(a) 5/15 (b) 1/9 (c) 5/15 x 4/14 (d) 2/9 (e) 5/21

CHAPTER 4

Opportunity Loss and Cost of Uncertainty

4.1. Tables and Definitions

4.1.1 Pay Off Tables

A payoff table indicates a conditional profit (or negative profit) for every act given various events. Each cell in the table is a net amount of gain or loss for each act-event combination.

To construct a pay-off table:

1) List every act-event combination with the acts across the top and the events down the side or vice versa.
2) For each act-event combination (each cell in the table) subtract the outflows from the inflows to get the net amount.
3) Minus signs must be preserved and indicate negative profit.

4.1.2 Opportunity Loss-is the difference between the profit actually realized in a situation and the profit which would have been realized if the best possible decision had been made for the event which actually occurred. An opportunity loss can occur even though a profit is made.

4.1.3 Loss Tables-a loss table indicates the conditional loss for every act given various events. It can be

constructed directly in a manner similar to payoff tables or derived from a payoff table. To construct a loss table from a payoff table:

1) Indicate the greatest possible profit for each event by some notation such as a star or circle around the largest profit in each row of the payoff table.
2) To obtain each entry in the loss table subtract the corresponding entry in the payoff table from the starred or circled entry in its row.
3) Note algebraic signs in the payoff table and that all cells in the loss table will be positive.

4.1.4 Expected profits and losses

1) The expected profits and losses from all acts can be found by taking a weighted average of each act (column) in the payoff and loss tables. The weights are probabilities assigned to each event.
2) The least expected loss and largest expected profit will both occur with the same act.
3) The difference between expected profits and losses of any two acts is equal, the profit being less and the cost being more by the same amount.

4.1.5 Cost of Uncertainty-The expected loss of the best decision.

4.1.6 Cost of Irrationality-The amount by which the expected

loss of a chosen decision exceeds that of the best decision.

Problems

1. The Navy buys an electron tube for $7.50. It costs the Navy $1.50 to pack and ship each tube issued. If a tube is not issued within one year, it no longer meets Navy specifications and must be disposed of.

   (a) If unissued tubes can be sold to commercial users for $5.00, how many tubes should the Navy stock at the beginning of each fiscal year? Assume that the Electronic Supply Office is operating under the Industrial Fund and makes $.25 on each tube issued. The following experience data is available on the probability of demand.

Demand (in 1,000's)	Probability
25	.03
30	.05
35	.15
40	.60
45	.10
50	.07

Solution:

Construct a payoff table and compute expected profits.

Demand	Proba-bility	Stock 25 Cond.	Stock 25 Exp.	Stock 30 Cond.	Stock 30 Exp.	Stock 35 Cond.	Stock 35 Exp.
25	.03	6.25	.1875	6.25	-.1875	-18.75	-.5625
30	.05	6.25	.3125	7.50	.3750	- 5.00	-.2500
35	.15	6.25	.9375	7.50	1.1250	8.75	1.3125
40	.60	6.25	3.7500	7.50	4.5000	8.75	5.2500
45	.10	6.25	.6250	7.50	.7500	8.75	.8750
50	.07	6.25	.4375	7.50	.5250	8.75	.6125
			6.2500		7.0875		7.2375

Problem 1 Pay off table (cont'd)

Demand	Proba-bility	Stock 40 Cond.	Stock 40 Exp.	Stock 45 Cond.	Stock 45 Exp.	Stock 50 Cond.	Stock 50 Exp.
25	.03	-31.25	-.9375	-43.7500	-1.3125	-56.25	-1.6875
30	.05	-17.50	-.8750	-30.0000	-1.5000	-42.50	-2.1250
35	.15	- 3.75	-.5625	-16.2500	-2.4375	-28.75	-4.3125
40	.60	10.00	6.0000	- 2.5000	-1.5000	-15.00	-9.0000
45	.10	10.00	1.0000	11.2500	1.1250	- .75	- .0750
50	.07	10.00	.7000	11.2500	.7875	12.50	.8750
			5.3250		4.8375		-16.3250

Using expected profit as a decision criterion, the Electronic Supply office should stock 35,000 tubes.

b. What is the cost of irrationality if the decision is made to stock 40,000?

(7.2375 - 5.3250) x 1000 = $1912.50

c. What is the cost of uncertainty?

Solution:

Construct a loss table and compute expected losses. Short-cut: Since the cost of uncertainty is the expected loss of the best decision, and the best decision is to stock 35,000, then expected loss need be computed for this stocking level only.

Answer: $2637.50

CHAPTER 5

Incremental Analysis

5.1 The incremental analysis of the problem stated on p. 68 of Schlaifer is given below on the basis of incremental profit.

Table 5.1 (From Schlaifer)

z	P(z)	$P(\tilde{z} < z)$	$P(\tilde{z} \geq z)$
ø	.05	.00	1.00
1	.10	.05	.95
2	.25	.15	.85
3	.30	.40	.60
4	.20	.70	.30
5	.10	.90	.10
6	.00	1.00	.00

PAYOFF TABLES

Stock 1	Cond. Value Stock	Do Not Stock	Incremental Expected Profit
$P(\tilde{z} < 1) = .05$	-$2	$0	-.10
$P(\tilde{z} \geq 1) = .95$	1	0	.95
			$.85

Stock 2	Cond. Value Stock	Do Not Stock	Incremental Expected Profit
$P(\tilde{z} < 2) = .15$	-$2	0	-$.30
$P(\tilde{z} \geq 2) = .85$	1	0	.85
			$.55

From our previous discussion we know that the probabilities of all possible events must add up to 1.00. Careful examination of table 5-1 reveals the same and is stated here:

that $$P(\tilde{z} \geq j) = 1.00 - P(\tilde{z} < j) \qquad (2)$$

Substitute in equation (1):

$$k_p[1 - P(\tilde{z} < j)] \geq k_n P(\tilde{z} < j) \qquad (3)$$

then

$$k_p - k_p P(\tilde{z} < j) \geq k_n P(\tilde{z} < j) \qquad (4)$$

$$k_p \geq P(\tilde{z} < j)(k_p + k_n) \qquad (5)$$

and finally:

$$P(\tilde{z} < j) \leq \frac{k_p}{k_p + k_n} \qquad (6)$$

Now all that we have to do to determine the best number to stock is compute the cumulative probabilities for $\tilde{z}$, demand less than the jth unit, and the critical ratio and compare. Select the largest j where the probability of it, is just less than the critical ratio.

Then

j	$1P($\tilde{z} \geq j$)	-	$2P($\tilde{z} < j$)	=	Expected Incremental Profit	Cumulative Sum ΣIP
1	.95		- .10		.85	.85
2	.85		- .30		.55	1.40 *
3	.60		- .80		- .20	1.20
4	.30		-1.40		-1.10	.10
5	.10		-1.80		-1.70	-1.60
6	.00		-2.00		-2.00	-3.60

* The best decision is to stock 2.

Selection of the best number to stock is based on the j which gives the greatest total expected profit. When the total expected profit is diminished by stocking one more, that one is one too many. Let's find an easy way to make this determination:

Let k_p = the positive incremental profit resulting from stocking the jth unit if sold.

k_n = the negative incremental profit resulting from stocking the jth unit if not sold.

We want the product obtained by multiplying the incremental profit by the cumulative probability of demand for the jth unit to be greater than the product obtained by multiplying the negative profit by the cumulative probability of no demand for the jth unit and express this relationship as follows:

$$k_p P(\tilde{z} = j) > k_n P(\tilde{z} < j) \qquad (1)$$

5.3 Problems

1. A retailer sells an item for $4. The cost of the item is $1.50. Each item sold is delivered to the customer, the cost of delivery being $.50 per item. If an item is not sold at the end of the day, it is disposed of at a salvage value of $.30. The retailer assigns probabilities to demand for the item in accordance with the following schedule:

Demand	Probability
less than 5	0
6	.05
7	.10
8	.12
9	.0
10	.20
11	.15
12	.10
13	.06
14	.02

a. Determine the optimum stocking level.

b. At the optimum level, compute the expected profit and expected loss.

c. What is the cost of uncertainty?

2. As the junior officer in the crew of a Navy transport aircraft, you are responsible for ordering an adequate supply of flight rations for the crew and passengers for each flight. Your next flight is a "R & R" run from Ugga Buggs Airport to Hong Kong, transporting military personnel on leave from the ships tied up in the Ugga harbor.

The Ugga Bugga municipal laws require that all flight rations for aircraft operating from the Ugga Bugga Airport be purchased from the Airport Cafe (operated by the Mayor's brother). The Cafe has established the following very firm rules on box lunch orders:

1. Box lunches will be ordered and paid for 48 hours before departure time. (price $1.50 each.)
2. If your order is larger than what you actually need at departure time (based on actual number of crew and passengers aboard) they will refund $ .50 to you for each lunch not required.
3. If you need additional lunches (your order was less than the number of passengers and crew aboard) they can be purchased at the cafe for $3.00 each right up to departure time.

A firm list of the total number of crew and passengers

for the flight will not be available until 6 hours before the departure time. Each passenger and crew member aboard will pay you $1.50 for their box lunch; however, any extra cost caused by excessive or late ordering will come out of your pocket.

After considerable discussion among the officers of the crew, the following probabilities of aircraft load (passengers and crew) were estimated based on the number of ships in port, how long it has been since pay day, etc.

Load	Prob.	Load	Prob.	Load	Prob.
less than 75	0	81	.07	88	.06
75	.01	82	.08	89	.05
76	.02	83	.09	90	.04
77	.03	84	.10	91	.03
78	.04	85	.09	92	.02
79	.05	86	.08	93	.01
80	.06	87	.07	more than 93	.00

How many box lunches should you order?

3. The U. S. NAVY owns 3 Polaris submarines. (As can readily be seen, the data in this problem is unclassified and may be revealed to the general public.) The Navy also has available only 24 skilled technicians to repair these three subs (which are constantly in need of repair). The subs, because of their important mission, can be in port only 8 hours each day (the home ports are not the same for the three subs). When the daily repairs are such that the Navy technicians cannot complete them within eight hours, (this can be determined at the beginning of the

work-period) additional civilian technicians areuutilized. The excess cost per day per civilian technician amounts to $10 more than the out-of-pocket cost of employing one of the Navy technicians. The total number of technicians (Navy or Navy and Civilian as the case may be) required for the past 100 days is shown in the table below.

Number of Technicians required	Number of Days		
	Sub A	Sub B	Sub C
0	0	0	4
1	2	0	4
2	6	2	6
3	10	4	10
4	10	4	12
5	12	8	15
6	12	10	15
7	10	12	10
8	8	12	12
9	8	15	3
10	7	15	3
11	8	10	3
12	7	8	3
	100	100	100

a. Compute the optimum assignment of the Navy technicians to the three subs.

b. Under this optimum assignment, what is the daily expected excess cost of hiring civilian technicians at each sub.

CHAPTER 6

Measures of Location:
Fractiles and Expectations;
Linear Profits and Costs

6.1. Fractiles

6.1.1 A fractile is a measure of location in a distribution. The "point f" fractile is the value in a distribution that is both:

1) Equal to or greater than a fraction (.f) of the values in the distribution (see a below) and
2) Equal to or less than a fraction (1-.f) of the values in the distribution (see b below)

Additionally:

1) The same value in a distribution can be included in both the lower (.f) part and the upper (1-.f) part (see c below) or
2) A fractile can include a range of values in a distribution (see d below)

To illustrate the above with a graphical example assume the following distribution of Lieutenants in the Navy by age group: $F_{.3} = 25$

$F_{.3} = 25$ a) This means that .3 of the distribution of Lieutenants in the Navy by age group are 26 years old or less. To state it another way, this fractile (.3) is equal to or greater than the fraction (.3) of the values (23, 24, 25, and 26)

in the distribution.

b) This also means that .3 of the distribution includes values that are equal to or less than a fraction (1-.f) or .7 of the values in the distribution. This is obviously true in the example since the number of Lieutenants who are 26 years old or less is less than the number who are 28 years old or less.
This analogy makes more sense when applied to the median which is a special case of a fractile; namely, the .5 fractile

$F_{.45\ (lower)} = F_{.55\ (upper)} = 27$

c) This is an example of the fact that the same value (27) in a distribution can be included in both the lower part (.f) and the upper part (1-.f) of a distribution. Therefore in our example we may think of some of the Lieutenants in the Navy who are 27 years old or less as belonging both in the lower .45 of the distribution and in the upper .55 of the distribution.

Table 6.1 Hypothetical Distribution of Lieutenants in the Navy by Age Groups

Age Group $\tilde{Z}$	Number in Group	Relative Frequency $\tilde{Z}$	Cumulative Frequency of Z or less
< 23	0	0	0
23	750	.05	.05
24	750	.05	.10
25	1,500	.10	.20
26	2,250	.15	.35
27	3,750	.25	.60
28	3,000	.20	.80
29	1,500	.10	.90
30	750	.05	.95
31	750	.05	1.00
> 31	0	0	
	15,000	1.00	

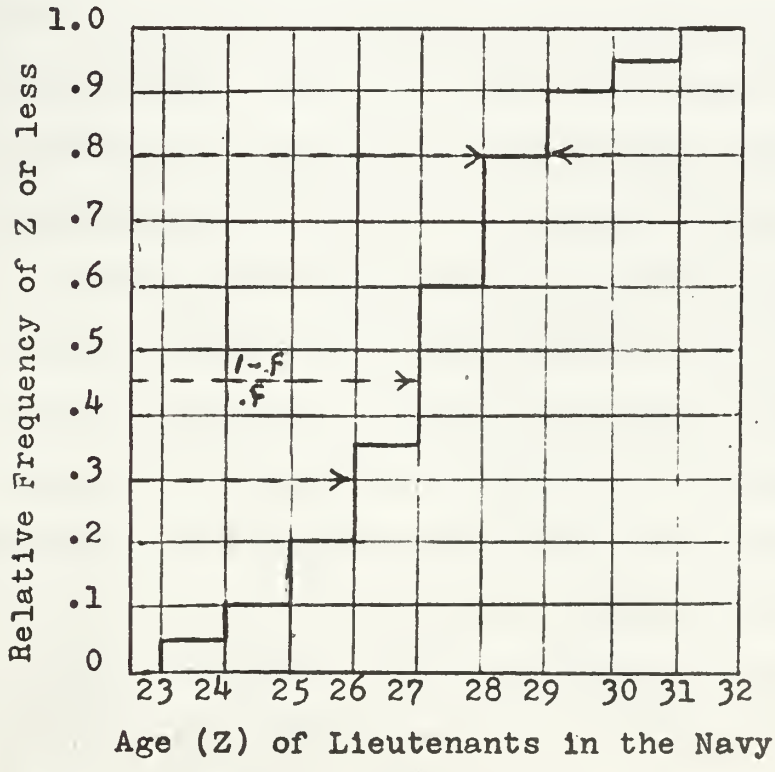


Figure 6.1

$F_{.8}$ = 28-29 inc.

d) This is an example of the fact that a fractile can include a range of values in a distribution and occurs in the "flat" areas of the graph rather than the vertical areas. In contrast to (c) above it means that all the Lieutenants 28 years old or less belong in the lower .8 whereas all those of 29 or more years belong in the upper .2 of the distribution.

6.1.2 Fractiles are also sometimes expressed as quartiles, deciles or percentiles. This is the result of dividing the total frequency into equal parts; four, ten or one hundred, respectively. Any two or more such positional measures can reproduce the frequency distribution. Thus, 40% of the observations lie between the median and the ninetieth centile. The second quartile is the same as the median. Note the difference between the first quarter, say, and the first quartile; the first quarter is a group of observations or a range of the variable, while the first quartile is the boundary point between the first and second quarters. An observation can be within the highest tenth, but not within the highest decile. The expression-"in the Xth centile"- is common but inaccurate.[1]

[1]Allen W. Wallis and Harry V. Roberts, Statistics; a New Approach, (New York, The Free Press of Glencoe Inc., 1962) p. 249

6.1.3 The median, or .5 fractile, is the middle observation. But to be precise, a certain amount of hair-splitting is necessary. It is a number in a set of observations that neither exceeds nor is exceeded by more than half the observations. When there is an odd number of observations, the median is the middle observation; when there is an even number of observations, it can be any number at or between the two middle observations. Thus, it may be a value not in any of the observations.[2]

6.1.4 The arithmetic mean is simply an average and it can be found in three ways:

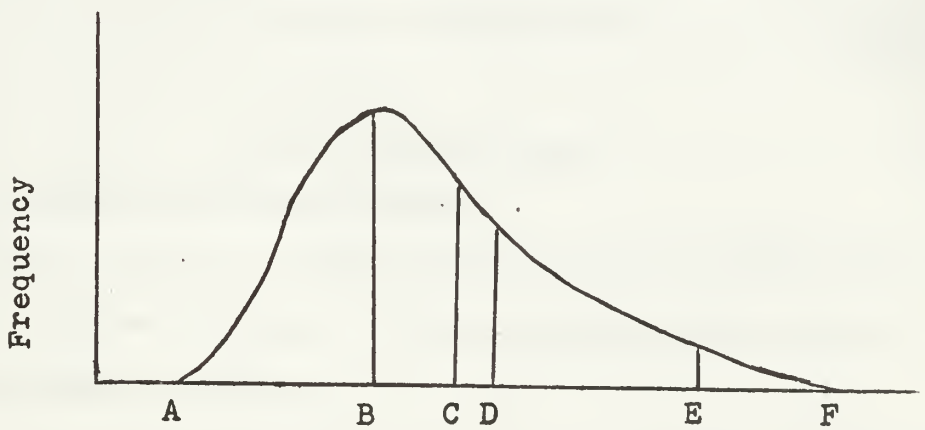
1) The most common way is to add the set of values and divide by the number of values comprising the set.
2) When the values of a random variable occur repetitively in a distribution, the arithmetic mean can be found by multiplying each value by the number of individual occurrences, summing the products and dividing the sum by the total number of occurrences. The number of individual occurrences may be considered as weights and the arithmetic mean as a weighted average.
3) The number of occurrences (or weights) can be converted to relative frequencies (whose sum is 1),

[2] Ibid. P. 216

multiplied by the corresponding value of the random variable and summed to get the mean.

6.1.5 The <u>mode</u> is the value in a set of observations that occurs most often. Of the three measures of location in a frequency distribution, the mean and the median are the principal ones, but the mode also has its useful purposes.

6.1.6 Graphical examples of the above definitions are illustrated in the following diagram with thumbnail definitions:



Variable (Z)

Figure 6.2

The distribution of the random variable (Z) has the following locations:

A-the minimum value

B-mode, the value of the variable (Z) which has the greatest number of observations

C-median, the value of the variable which exceeds half of the observations and is exceeded by half

D-mean, the value ordinarily referred to as "average"

E-the ninetieth centile, the value of the random variable (Z) which exceeds 90 percent of the observations and is exceeded by 10 percent.

F-the maximan value[3]

6.2 Linear Profits and Costs

6.2.1 When we compute the mean of a probability distribution of a random variable, we are computing the expected value of the random variable.

In many cases profit (or cost) is just a special case of a random variable and is the mean of a probability distribution of conditional profits (or costs). This can be expressed in the form:

$$\text{Conditional profit or cost} = K + k\,E(\tilde{Z}),$$

[3] Ibid. pp. 211-212

Where K and k are constants and E $(\tilde{Z})$ is the expected value or mean of the distribution of a random variable $\tilde{Z}$.

Problems

1. The Navy formerly ran its own coffee roasting and distribution plant. Assume that the table below is the monthly demand of the various ships in the Navy for coffee. Also assume that the plant pays $1.00 per pound for green coffee, that grinding and distribution costs are $.25 per pound sold and that a profit of $.25 per pound is authorized as a contribution to the WAVES Retirement Fund. If the Navy plant runs short, the demand is met by purchasing coffee from commercial sources at $1.50 per pound including shipping costs to the requisitioning activity.

Demand in lbs	No. of ships	Demand in lbs	No. of ships
50	21	110	114
60	51	120	69
70	53	130	45
80	81	140	21
90	120	150	15
100	150		

a) Considering the above data as a probability distribution rather than a frequency distribution, graph $P(\tilde{Z} \leq Z)$ against the demand Z.

b) Compute the deciles ($F_{.1}, F_{.2} \ldots F_{1.0}$).

What is the value $F_{.8}$? Answer: 110 to 120 lbs. per ship.

c) What is the expected monthly demand per ship?

Answer: 96.88 lbs.

d) In anticipation of reduced demand for coffee during the summer months the plant maintains a total stock of 60,000 lbs. What is the monthly expected loss to the Retirement Fund if demand falls to an average of 85 lbs. per ship?

Answer: $937.50

e) If ordering, receiving and processing green coffee from Brazil takes one month half the time, two months a third of the time, and three months the remainder of the time, how much coffee should be maintained in stock to meet the normal expected demand?

Answer: 121,000 lbs.

f) Assume that there is no longer a requirement to contribute to the WAVE Retirement Fund but the plant is still authorized to make $.25 per pound profit and reduce inventory by using the proceeds to purchase coffee through commercial sources. How much coffee should the Navy plant process and how much should be obtained commercially in order to just break even?

Answer: 62,281 lbs. and 10,379 lbs. respectively.

g) Assume that coffee not sold to a ship by the end of the month is no longer consumable and must be surveyed at a loss of $1.00 per pound, its purchase price. If the plant manager uses the frequency distribution in the above table, how many pounds per ship per month should be stocked?

Answer: 80 lbs.

CHAPTER 7

Assessment of probabilities by historical smoothing.

7.1 Relating probabilities to relative frequencies.

When occurrence of a number of events are converted into relative frequencies, it is possible to assess probabilities to each of these events based on this historical data. The historical data however, may generate a distribution that appears irregular and is difficult to manipulate. On such occasions, one might wish to adjust the assigned probabilities to approximate a smooth curve rather than assess probabilities from frequencies derived from the raw data provided that there are no assignable causes for the variations.

For example, take the following 106 time lapses (in minutes) between eruptions of Old Faithful Geyser in Yellowstone National Park. (they were recorded between the hours of 8 am and 10 pm on July 6- July 14, 1955, and they were measured from the beginning of one eruption to the beginning of the next).[1]

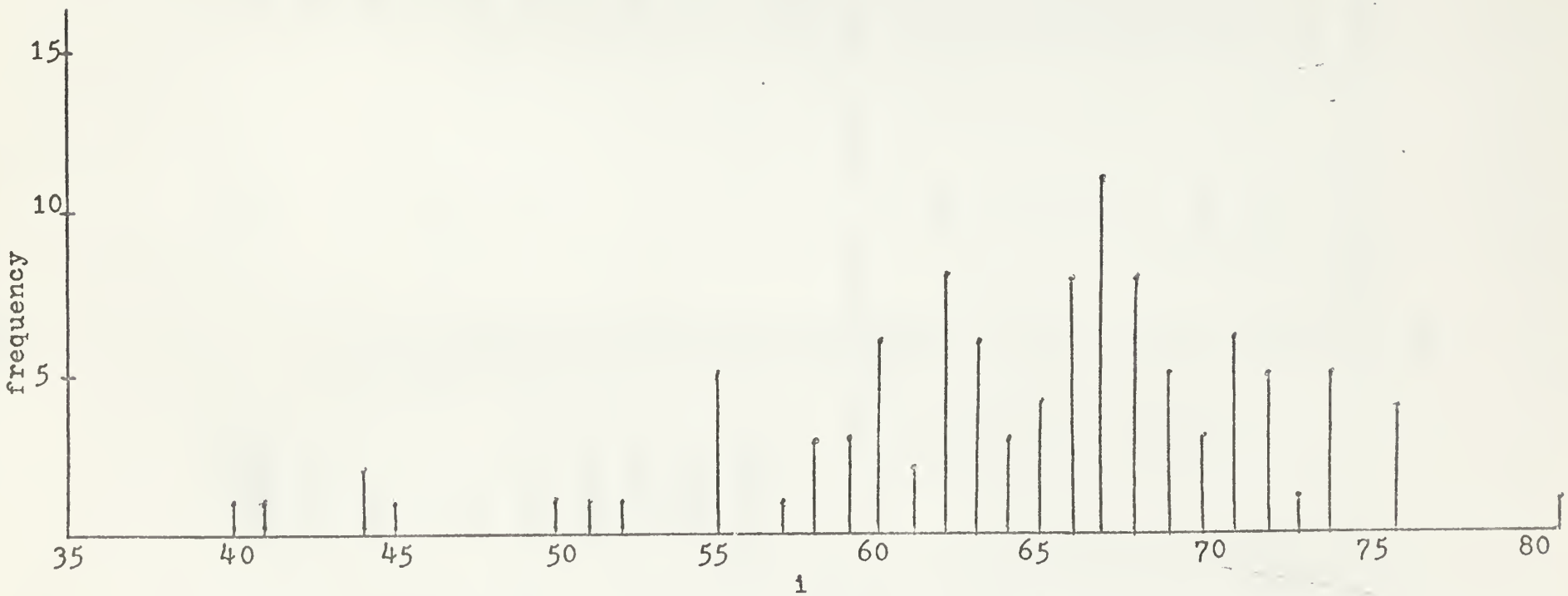
[1] John E. Freund, and Frank J. Williams, Modern Business Statistics, p. 31.

Table 7-1

60	62	68	72	50	76	59	69	62	70	67	44	71
63	62	65	69	62	70	40	62	74	59	68	66	68
66	55	60	65	63	57	66	68	67	58	74	74	66
63	62	61	66	67	67	74	76	63	40	63	71	67
61	67	71	64	52	76	63	71	76	60	74	41	72
44	73	60	69	55	81	55	68	51	67	62	66	66
60	72	68	59	69	55	58	60	69	58	66	63	68
62	55	67	67	64	45	65	60	64	67	71	72	72
41	67											

Table 7-2 is the data rearranged by frequency (f) vs. time intervals (i) and is plotted in figure 7-1.

From this data, one could assign probabilities to the occurrence of each event. However, some events would be assigned a probability of zero for example, $P(\tilde{i} = 47) = \emptyset$. This quite obviously is not a very good assessment of that probability. Suppose this date is reclassified into 5 minute intervals. Table 7-3 is generated on this premise.



Display of frequency v. time interval to the nearest minute between eruptions of Old Faithful Geyser at Yellowstone National Park.

Fig. 7-1

Table 7-2

Data from Old Faithful classified in 1 minute intervals.

i	f	i	f	i	f	i	f
40	1	55	5	63	6	70	3
41	1	57	1	64	3	71	6
44	2	58	3	65	4	72	5
45	1	59	3	66	8	73	1
50	1	60	6	67	11	74	5
51	1	61	2	68	8	76	4
52	1	62	8	69	5	81	1

Table 7-3

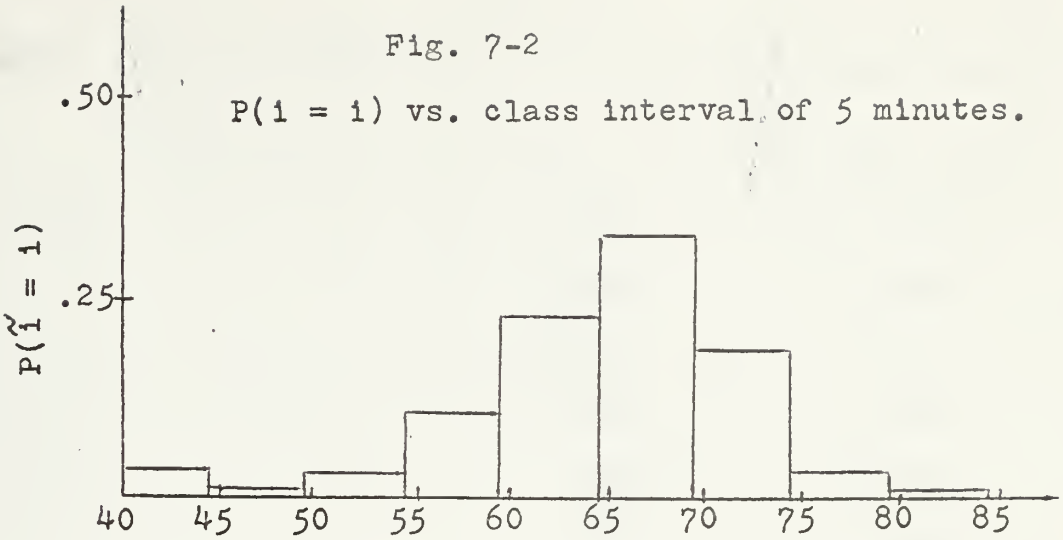
Data from Old Faithful classified in 5 minute intervals.

Table

i	f	$P(\tilde{i} = i)$	$P(\tilde{i} \geq i)$
40-44	4	.0378	1.000
45-49	1	.0094	.9622
50-54	3	.0283	.9528
55-59	12	.1133	.9245
60-64	25	.2360	.8112
65-69	36	.3390	.5752
70-74	20	.1890	.2362
75-49	4	.0378	.0472
80-84	1	.0094	.0094
	106	1.0000	

Fig. 7-2

P(i = i) vs. class interval of 5 minutes.



In figure 7-2, it is apparent that the distribution has smoothed out considerably, but there is still some disparity at the left end of the distribution. By manipulating the data, some more irregularities may be removed such as the bimodal aspect, simply by changing the interval and/or where each class begins. Suppose that the data is reclassified into ten minute intervals. Table 7-4 is generated on this premise.

Table 7-4

Data from Old Faithful classified in ten minute intervals.

i	f	P(i)	P(i = i)
31-40	1	.0094	1.0000
41-50	5	.0472	.9906
51-60	20	.1885	.9434
61-70	58	.5475	.7549
71-80	21	.1980	.2074
81-90	1	.0094	.0094
	106	1.0000	

Fig. 7-3

$P(\tilde{i} = i)$ vs. class interval of ten minutes.

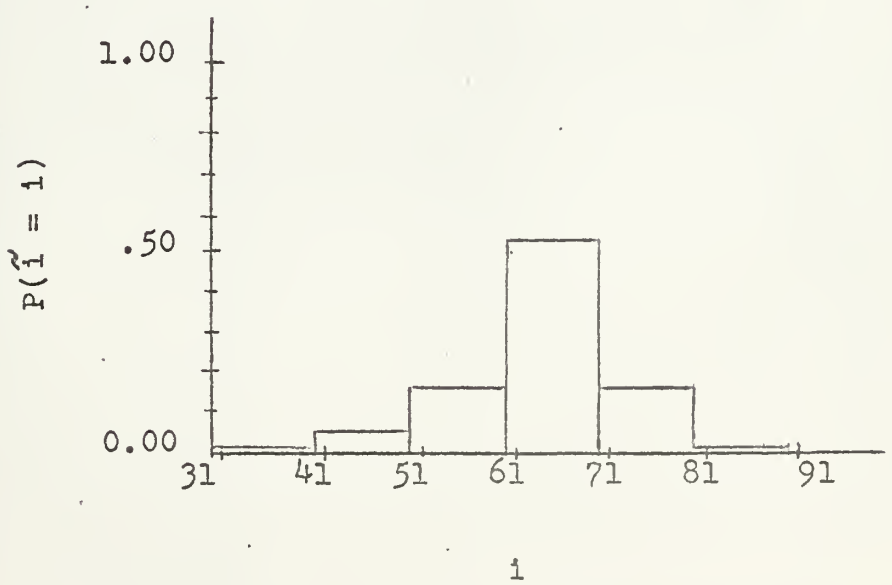
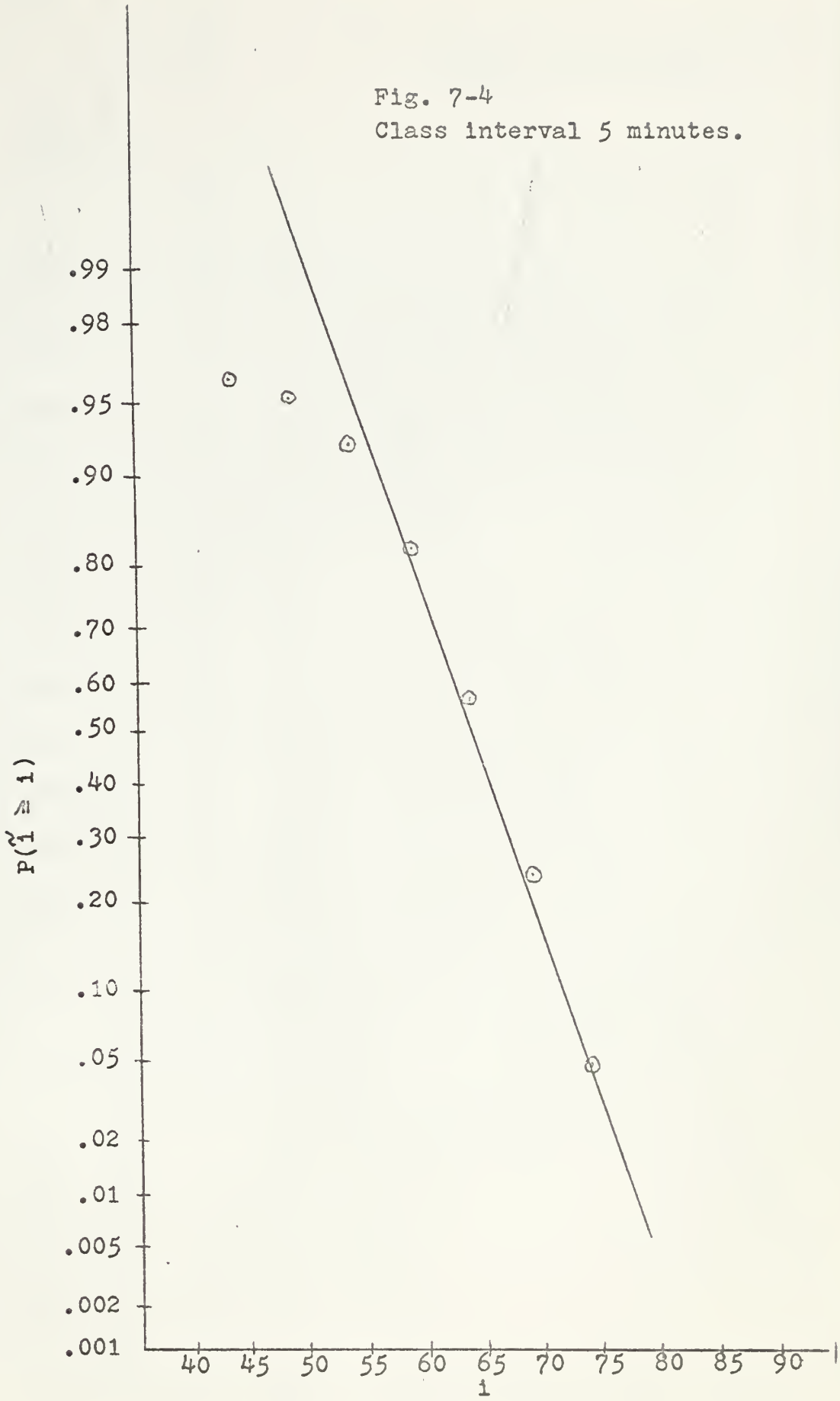


Fig. 7-4
Class interval 5 minutes.



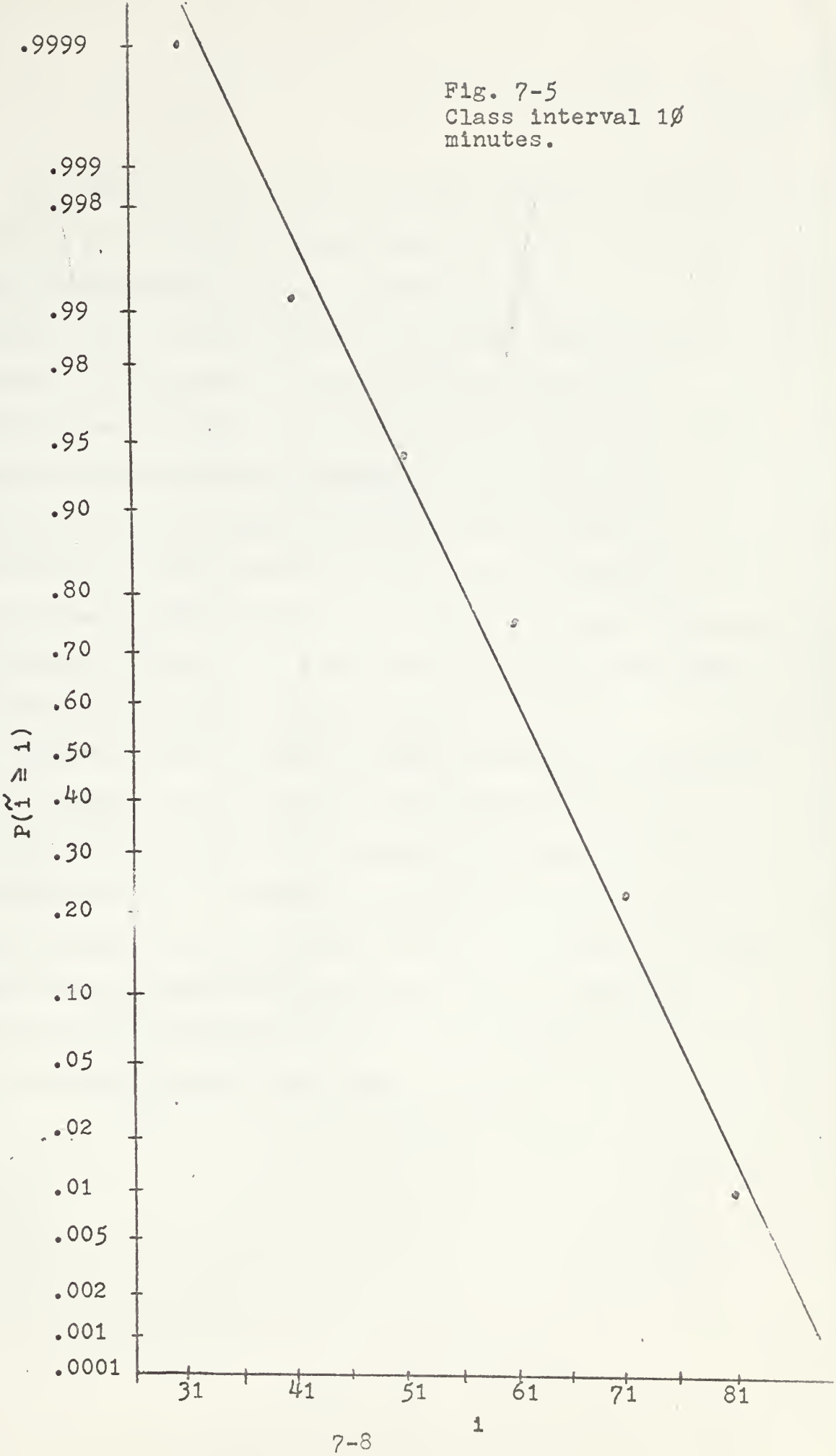


Fig. 7-5
Class interval 1Ø minutes.

In figure 7-3 is $P(\tilde{i} \geq i)$ plotted against a class interval of 10 minutes derived from table 7-4. Notice that classification intervals can begin and end at any selected point on the x-axis. By merely changing these points it is possible to make profound changes in a distribution depending on the diversity of data and classification intervals selected. It should also be emphasized that although tables 7-3 and 7-4 show probabilities to four places, they are <u>not</u> significant to four places. Additional decimal places have been arbitrarily filled in beyond the significant digits to complement all probabilities to a sum of 1.0000.

Figure 7-4 and 7-5 are plotted on normal probability paper. Figure 7-4 is a plot of the $P(\tilde{i} \geq i)$ based on a classification interval of 5 minutes and figure 7-5 is a plot of the $P(\tilde{i} \geq i)$ based on a classification interval of 10 minutes. As can be seen in the illustrations, proper selection of classification interval and the starting point are primary factors in the determination of a smooth curve derived from empirical data.

CHAPTER 8

Conditional, Joint, and Marginal Probability

8.1 Theory of Probability

8.1.1 The theory of probability is a branch of mathematics and one of the applications of this theory is statistics. Probability, chances and likelihood are words that are common in everyday conversation--the probability that it will rain tomorrow, the chances that the Yankees will win the Pennant this year, the likelihood that there is life on Mars, etc. However, in order to be useful, mathematical probability should be expressed quantitatively. Probabilities used in business may be based on past experience, a collection of data, sampling, or an estimate of the future. In the application of probability theory to business and industry, it is important not to confuse probability with frequency even though the two can often be equated (see section 3.2 of Schlaifer).

8.1.2 For our purposes, probability can be broken down into three concepts: joint, conditional and marginal. Basically:

a) Joint probability is the probability that two or more events will all occur; each is independent of the other.

b) Conditional probability is the probability that an event will occur on condition that another event has already occurred or would occur; the second event is dependent on the first; this definition also holds true for more

than two events.

c) Marginal probability of an event is the sum of the joint probabilities making up the event.

8.1.3 The above definitions can be illustrated by the following example:

Table 8-1 Age and Sex of Employees of Corporation

	Age	Male B_1	Male B_2	Total
A_1	< 34	2,100	900	3,000
A_2	34-54	4,200	1,800	6,000
A_3	≥ 55	700	300	1,000
		7,000	3,000	10,000

Joint Probability-The probability that an employee selected at random is of a specified sex and age is a joint probability because two characteristics of the random selection have been specified. The probability of selecting a male worker at random is $P(B_1) = \frac{7,000}{10,000} = .7$. The probability of selecting a male worker under 34 years old is the joint probability $P(A_1B_1) = \frac{2,100}{10,000} = .21$

Conditional Probability-The probability of selecting an employee less than 34 years old when it is known that he is male is a conditional probability because one event is already known, namely, the condition that he is male.

Then, $P(A_1|B_1) = \frac{2,100}{7,000} = .3$

Marginal Probability-The marginal probabilities appear in the summation column and row of the above table, thus the term "marginal". The marginal distribution of sex of employees is: male = .7 and female = .3. The marginal distribution by age groups are: $<34 = .3$, $34-54 = .6$ and $\geq 55 = .1$. Note that the sum of each margin is 1.

The above explanation may be made clearer by restating Table 8.1 in the usual probability function:

Table 8.2 Frequency Distribution of Employees by Age and Sex

	Age	Male B_1	Female B_2	Total
A_1	< 34	.21	.09	.30
A_2	34-54	.42	.18	.60
A_3	≥ 55	.07	.03	.10
		.70	.30	1.00

If we want to know the probability that a person selected at random is either male or under 34 years old, we can add the probabilities of each, but we must be careful not to add the same group twice. Thus the probability that a person is male is .7 and the probability that a person is under 34 years old is .3 but we have counted the male under 34 twice. Therefore the answer is .7 + .3 - .21 = .79 The answer could also have been obtained in several other ways, e.g.; the probability that a person is male is .7,

but to include all the people under 34 we must also include females, .7 + .09 = .79.

Try similar reasoning for .3 + .42 + .07 = .79

Problems

1. If 70% of collisions at sea occur during reduced visibility and 60% occur in the approaches to major world harbors,

   a) What per cent occur in the open sea in good visibility?

   b) In reduced visibility in the open sea?

   Answers: a) 12%, b) 28%

2. In the above problem given that a collision occured in the open sea, what is the probability that visibility was good? Answer: .3

3. In the destroyer force a ship can earn a permanent gold "E" if she wins the Battle Efficiency Award five years in a row. What is the probability that a relatively new ship which won no awards her first two years in commission then won three consecutive awards can win the gold "E"? Assume that there are eight ships in the squadron and (although unrealistic) they are evenly matched and therefore, have equal opportunity to win.

   Answer: .0156

4. If 20% of the men in the Navy are not married and 70% of those who are have two or more children while

30% have three or more, what per cent are married with one or no children?

Answer: 24%

5. Seventy per cent of the time wholesale and retail prices move in the same direction. Sixty per cent of the time when they move together they are moving upward, and seventy five per cent of the time when they move upward together a prosperous period follows; twenty five per cent of the time bad times follow. There are as many prosperous periods as bad periods. When prices move in opposite directions, the odds are even that the following period will be prosperous or bad. What is the probability that there will be a downward movement in both prices followed by a prosperous period?[1]

Answer: .035

Hint: Try to get the answer by filling in the rest of this table:

		Same	.70	Opposite
		up	down	
Prosperity	.50			
Bad Times				

1 Perlman, Richard, Problems in Statistics for Economics and Business Students. New York: Holt Rinehart and Winston, Inc., 1963. P. 27

6. Three destroyers in a ASW exercise have a past record of making contact 20, 30, and 40 per cent of the time, respectively. Assuming that there is a different submarine in the vicinity of each destroyer, (a) what is the probability that all make contact? (b) that at least two will make contact? (c) that none make contact?

   Answers: a) .024, b) .212, c) .336

7. A cruiser is attacked by four airplanes in succession. If her past performance has shown an ability to shoot down 40% of attacking aircraft, (a) what is the probability of shooting down exactly one plane? (b) of shooting down all four? (c) of shooting down at least two?

   Answers: a) .3456, b) .0256, c) .5248

8. A dealer holds 2 aces and 4 dueces in his hand. He offers to form a pool of $60 to be paid to the first man among five to draw an ace. If A, B, C, D, and E are going to draw in that order, how much should each pay to play? A card once drawn is not replaced.

   Answers: A-$20, B-$16, C-$12, D-$8, E-$4

9. Comment on the following quotation:

   . . . .Nothing . . .is more difficult than to convince the merely general reader that the fact of sixes having been thrown twice in succession by a

player at dice, is sufficient cause for betting the largest odds that sixes will not be thrown in the third attempt. A suggestion to this effect is usually rejected by the intellect at once. It does not appear that the two throws which have been completed, and which lie now absolutely in the Past, can have influence upon the throw which exists in the Future. The chance for throwing sixes seems to be precisely as it was at any ordinary time . . . .And this is a reflection which appears so exceedingly obvious that attempts to controvert it are received more frequently with a derisive smile than with anything like respectful attention. The error here involved--a gross error redolent of mischief--I cannot pretend to expose within the limits assigned me at present; and with the philosophical it needs no exposure . . . . .[1]

[1] Edgar Allan Poe, "The Mystery of Marie Roget", final paragraph in Murders in the Rue Morgue.

CHAPTER 9

The Binomial Distribution

9.1 Introduction.

In chapter 3, a binomial distribution was developed using a ten-sided die as a model and if you worked the first problem at the end of that chapter, you developed one yourself. The binomial probability is written in Schlaifer as $P(r) = C_r^n\ p^r\ q^{n-r}$. It is an expression for any term of the binomial expansion. For example, let's determine the fourth term of an expansion of the form $(p + q)^7$. It can be expressed $P(4) = C_4^7\ p^4\ q^3$. By performing the indicated operation we have

$$P(4) = \frac{7!}{4!3!}\ p^4\ q^3$$

$$P(4) = 35\ p^4\ q^3$$

By the method used in chapter 3, the 4th term coefficient from Pascal's triangle is 35 and the binomial is $p^4\ q^3$, so again the 4th term is $35\ p^4\ q^3$.

Any binomial distribution can be developed in the same manner. It is important to remember the relationship between p and q. The assignment of a value to either p or q implicitly determines the other since p is the 1's complement of q and vice versa, i.e., $p = 1 - q$. With this in mind it is easy to see that there are two parameters that determine a binomial distribution, n and p. In the dice model, p is the probability that is assigned to a

particular face coming up on the die, q is the probability of the same face not coming up, n is the number of throws or trials in a series, and r is the number of times that particular face comes up. It is sometimes more convenient to think of r in terms of the number of successes. Thinking of the die model is one way of keeping the parameters of the binomial distribution in perspective.

In the next section it will be shown how the binomial distribution is developed and conveniently structured into tables thus making it unnecessary to calculate each distribution.

9.2 Derivation of a specific Cumulative Binomial Distribution.

A key to the use and application of Table I in Schlaifer, is an understanding of how the table values were derived, and the organization of the data. To illustrate a derivation of part of the tables, let's derive

$$P(\tilde{r} \geq r | n=9,\ p=.40).$$

In the left hand column of Table 9-1 is the binomial distribution of $P(\tilde{r} \geq r | n=9,\ p=.4)$ reproduced from Table I in Schlaifer. To the right of the vertical line in Table 9-1, are the calculations which derive this same distribution to nine places. These calculations were made of this particular distribution because the probabilities are finite in nine places and it again emphatically demonstrates that the sum of the probabilities of all events in a distribution equal 1.000000000.

From Schlaifer

$n = 9$

$p = .40$

r	$P(\tilde{r} \geq r)$	$C_r^n \times p^r \times q^{n-r} = P(\tilde{r} = r \mid n=9, p=.40)$	$P(\tilde{r} \geq r \mid n=9, p=.40)$
0	1.0000	$1 \times (4/10)^0(6/10)^9 = .010077696$	1.000000000
1	.9899	$9 \times (4/10)^1(6/10)^8 = .060466176$	.989922304
2	.9295	$36 \times (4/10)^2(6/10)^7 = .161243136$	.929456128
3	.7682	$84 \times (4/10)^3(6/10)^6 = .250822656$	.768212992
4	.5174	$126 \times (4/10)^4(6/10)^5 = .250822656$	.517390336
5	.2666	$126 \times (4/10)^5(6/10)^4 = .167215104$	.266567680
6	.0994	$84 \times (4/10)^6(6/10)^3 = .074317824$	.099352576
7	.0250	$36 \times (4/10)^7(6/10)^2 = .021233664$	.025034752
8	.0038	$9 \times (4/10)^8(6/10)^1 = .003538944$	.003801088
9	.0003	$1 \times (4/10)^9(6/10)^0 = .000262144$	.000262144
		1.000000000	

Table 9-1

Table I in Schlaifer rounds off the probabilities to the 4th place. By rounding the right hand column in Table 9-1 to four places, the calculated quantities agree with Schlaifer. Notice that Table I in Schlaifer provides distributions for parameter n of values 1 through 20, 50, and 100, and for parameter p, of values .01 to .50 inclusive. The tables for values of p from .51 to .99 can easily be derived from Table I. The next section will show how this can be done.

9.3 Obtaining Cumulative Binomial Probabilities from Table I in Schlaifer when $p > .50$.

Examine the formula by which the binomial distribution was developed in Table 9-1:

$$P(\tilde{r} = r|n=9,\ p=.40) = C_r^9\ .40^r\ (1.00 - .40)^{9-r} \qquad (1)$$

$$= C_r^9\ .40^r\ (.60)^{9-r} \qquad (2)$$

For parameters n = 9, p = .60, the formula is

$$P(\tilde{r} = r|n=9,\ p=.60) = C_r^9\ .60^r\ (1.00 - .40)^{9-r}. \qquad (3)$$

$$= C_r^9\ .60^r\ (.40)^{9-r} \qquad (4)$$

$$= C_r^9\ .40^{9-r}\ (.60)^r \qquad (5)$$

By comparing formulas (1) and (5), one can see that

a) the coefficients are the same,

b) the quantities calculated from (2) are also calculated from (5) except in the inverse order,

c) and it follows that

of 1, 2, 3, and 4 = $P(\tilde{r} < 5 | n=9, p=.4)$,

c) in the 3rd quadrant, the sum of the probabilities of 9, 8, 7, 6, and 5 = $P(\tilde{r} \geq 5 | n=9, p=.6)$,

d) in the 4th quadrant, the sum of the probabilities of 4, 3, 2, 1, and Ø = $P(r < 5 | n=9, p=.6)$.

The sum of the probabilities in the 1st and 4th quadrants are equal, and the sum of the probabilities in the 2nd and 3rd quadrants are equal. Table I in Schlaifer gives the sum of the probabilities in the 1st quadrant. Since we know that the sum of the probabilities of events Ø to 9 is equal to 1.0000, then it follows that the sum of the probabilities of the 2nd quadrant is equal to 1.0000 minus the sum of the probabilities in the 1st quadrant. Thus, the sum of the probabilities in the 3rd and 4th quadrants are defined.

$P(\tilde{r} = \emptyset \mid n=9,\ p=.40) = P(\tilde{r} = 9 \mid n=9,\ p=.60)$

$P(\tilde{r} = 1 \mid n=9,\ p=.40) = P(\tilde{r} = 9 \mid n=9,\ p=.60)$

$P(\tilde{r} = 2 \mid n=9,\ p=.40) = P(\tilde{r} = 9 \mid n=9,\ p=.60)$

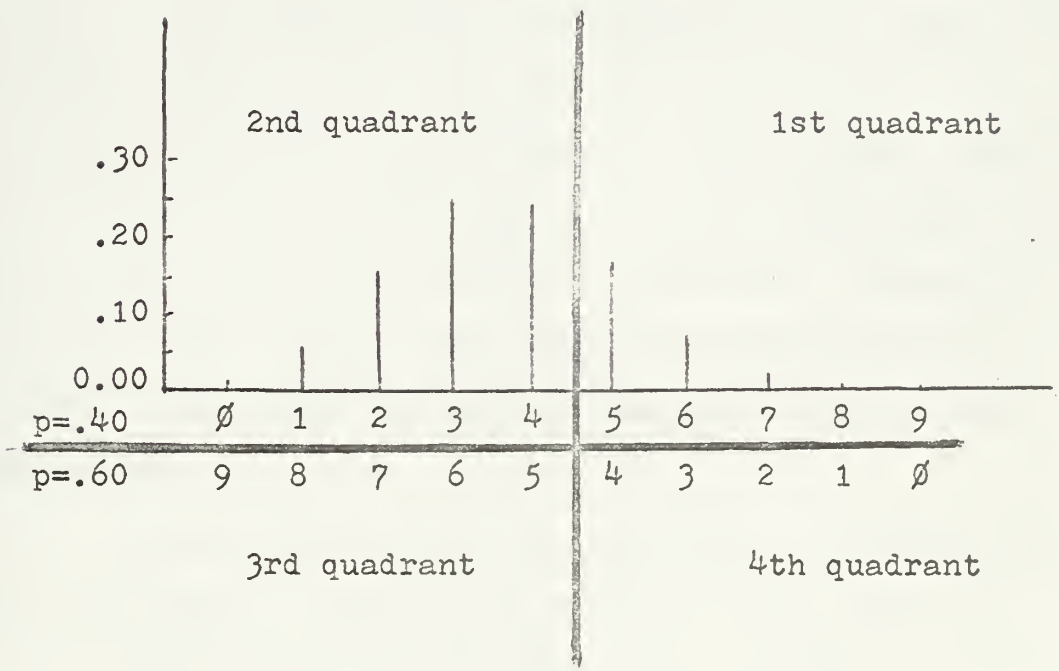
$P(\tilde{r} = 3 \mid n=9,\ p=.40) = P(\tilde{r} = 9 \mid n=9,\ p=.60)$

. . . etc.

A convenient way to illustrate (c) above is to draw a pictorial diagram as follows:

Fig. 9-1

A Pictorial Diagram.



In Fig. 9-1, the discrete binomial distribution is plotted to scale for the sake of explanation. From the diagram, it can be seen that

a) in the 1st quadrant, the sum of the probabilities for 5, 6, 7, 8, and 9 = $P(\tilde{r} \geq 5 \mid n=9,\ p=.40)$,

b) in the 2nd quadrant, the sum of the probabilities

Problems:

9-1. You are the commanding officer of a destroyer that is undergoing refresher training in San Diego. The ship is required to shoot an anti-aircraft gunnery exercise using VT non-frag ammunition which if not triggered by the target, is designed to self-destroy before water entry. In two practice exercise firings of 100 rounds per exercise, the ship has scored unsatisfactorily. The gunnery officer swears that the battery is properly aligned and that the cause of the unsatisfactory scores must be the ammunition. You know that the lots are supposed to be no more than 4 per cent defective. Your observers report that 7 rounds out of the first 100 did not self-destroy and that 9 rounds out of the second 100 did not self-destroy. Assuming a $p = .04$, what probabilities would you assign to these two events? What could you hypothesize about the per cent defective?

9-2. A navy supply center has 100 excess typewriters in various stages of repair. A notice is sent out to all ships in port that these typewriters will be issued on a first come, first served, no cost basis with a limit of 3 to destroyers; 9 to DLG's and 15 to cruisers. Some of these typewriters are beyond economical repair. No choice among the typewriters will be permitted. Assuming that all ships draw their limit:

CHAPTER 10

Statistical Decisions Rules and Their Error Characteristics

10.1 Terms

Decision makers must often make a choice between two alternatives such as-whether to publish or not; whether to market a new product or not; whether a new development (such as a drug) is effective or not; whether an industrial (or production) process is satisfactory or not; and in the social sciences whether certain data should be "believed" or not. Making a choice on matters of this type can be done by statistical analysis of random sampling, which leads to some new terms--Null Hypothesis, Alternate Hypothesis, Type I Error, Type II Error, Producer's Risk, Consumer's Risk and Tests of Significance.

10.2 The Null Hypothesis

The statistical hypothesis approach, in attempting to make a decision or reach an opinion, is to make an assumption for the sole purpose of either accepting it or rejecting it. For example, we make an assumption that a coin is fair: that is, the probability of getting heads on n tosses is $p = .5$. This we might call the null hypothesis. If we toss the coin 100 times and get 60 heads we might then reject the null hypothesis and accept the alternate, that the coin is not fair. Generally, the null hypothesis may be thought of as the status quo and leads to no action. However, this is not

necessarily true, for the hypothesis can be stated in a number of ways. The alternate hypothesis is the opposite of the null and leads to an opposite course of action.

10.3 Type I and Type II Error

Suppose that as a result of random sampling we are lead to a certain course of action. But what if the sampling, even though properly done, has produced untrue information? If we acted on this untrue information we have made an error. Returning to the example of the coin with a null hypothesis that it was fair but subsequent acceptance of the alternate based on a sample of 100 tosses. Suppose that we toss the same coin an additional 1000 times obtaining 510 heads and 490 tails, and further that by a scientifically accurate balancing test we find that the coin was in fact fair. Then, having previously rejected the null when in fact it was true, we have made an error. This is called a Type I error. A Type II error is just the opposite-to accept the null hypothesis when it is false-example: toss a coin 100 times with 48 heads and 52 tails leading to the null hypothesis that the coin is fair but subsequent tosses with 420 heads and 580 tails out of a thousand indicate that the coin is in fact not fair; then an error, Type II, has been made in accepting the null hypothesis when it was false.

This terminology may be illustrated by the following table:

Table 10.1

Reality	Decision	
	Accept the Null	Accept the Alt.
Null hyp. is true	No Error	Type I Error
Alt. hyp. is true	Type II Error	No Error

10.4 Producer's and Consumer's Risk

When testing any hypothesis by sampling techniques; the decision rule, in order to be useful, must minimize the probability of error. The difficulty here is that minimizing one type of error is usually accompanied by an increase in the other. For example, a drug company desires to test a new remedy on a sample of 100 patients in a hospital. A substantial amount of data over a period of years has lead to the conclusion that the old remedy is 50% effective. Therefore, if the new remedy results in 51% or more recoveries, it is better than the old and should be adopted. We set the null hypothesis that the new remedy is no more effective than the old and therefore the true proportion of recoveries will be .5 or less. If in the test we get 51 recoveries, we might assume that the null hypothesis is false and that the new remedy is better than the old. It is intuitively obvious that this improvement from 50 to 51 percent might very well have been by chance alone and there-

fore we might make a Type I error by rejecting the null hypothesis when in fact it was true. So in order to reduce the possibility that the test results are from chance alone, we might set the decision rule at, say, 56 recoveries. The probability of making a Type I error is now reduced: that is, if we state that in order for the new remedy to be considered better than the old there must be a 56% recovery as opposed to the old 50%, then we are more sure of our test and less likely to make a Type I error. However, we have now increased the probability of a Type II error; for, if in fact the new remedy is more effective than the old, we have, by increasing the level of the test it must meet, reduced the probability of finding this out. By this decision rule a remedy that was actually 54% effective might very well be rejected. It would be wrong to reject the new remedy, and by so doing a Type II error would be made. Thus it may be seen that in this example, as in many others, especially in marketing and quality control, the probability of making one type of error can only be reduced at the expense of increasing the probability of making the other type of error; both depend on the true value of the unknown p that is being tested. Errors of the type described above are also sometimes referred to as producer's and consumer's risk; the producer's risk is the probability that his improved product will be rejected and the consumer's risk is the

probability that an unimproved product will be accepted. Both types of errors can be reduced by taking larger samples and the nearer the sample size approaches the true population size, the more accurate will the decision rule prove to be. But the larger the sample size, the more expensive and time consuming the testing process.

10.5 Significance Level

Realizing that it might be unduly arbitrary to set a decision rule too rigorously, we might want to restate it by setting a maximum probability by which we would be willing to risk an error in our hypothesis; that is, we might be willing to accept a 5% probability of error; then we would be 95% confident that we have made the right decision. The maximum probability with which we would be willing to accept a Type I error is called the significance level and is denoted by α(Producer's risk). The maximum probability with which we would be willing to accept a Type II error is denoted by β (consumer's risk). Significance levels are most often set at .05 and .01, but other significance levels can be used and often are, particularly in medical and psychiatric statistics.

10.6 Decision Rules

Business problems of the type discussed in the basic text can be solved by first computing a breakeven point and stating the null hypothesis in terms of the breakeven point

and then setting up an opportunity loss table. Next a decision is made on just how large a risk the business (or producer) can take and how large the consumer can take. If it is desired to balance the two types of risks, the former α , and the latter, β , are set equal to each other. With this information and a set of binomial tables an (n,c) decision rule can then be found where n is the sample size and c is the critical (rejection or acceptance) number. If the two types of risks are not balanced, then the businessman (or producer) must decide the greatest chance of risk, α , that he is willing to take and then choose an (n,c) decision rule that meets this criterion while at the same time reducing the probability of Type II error as much as possible.

Problems

1. At a certain critical point in a production process a component must be tested at the manufacturer's expense. If the manufacturer will accept a risk of 5% of having what he considers a good lot (that is, 6% or less defective) rejected what rejection number should he choose from a sample of 50?

   Solution: given: $n = 50$, $\alpha = .05$, $p = .06$

   from the binomial tables: $R(6) - .0776$

   $R(7) - .0289$

   a rejection number of 7 would limit the manufacturer's risk to 2.89% but 6 would limit

it to 7.76%. Since the manufacturer does not want his risk to be any higher than 5%, he must choose the sampling plan (7,50).

2. Promotion material on a new drug claims that it is 90% effective.

   a) If this is <u>true</u>, what is the probability that in a sample of 100 there would be 11 non-effectives?

      Answer: .4168

   b) If the Null Hypothesis is that the new drug is no more effective than the old, how many non-effectives would have to turn up in the sample for it to be statistically significant at the .05 level?

      Answer: 16

   c) If 10 non-effectives were found in a sample what is its statistical significance?

3. In a Bernoulli process of sample size 20 and defectives $p = .25$, find the following:

   a) The probability of 6 or more defectives.

      Answer: .3828

   b) The probability of 6 or less defectives.

      Answer: .7858

   c) The probability of exactly 6 defectives.

      Answer: .1686

d) The probability of 5 to 8 defectives inclusive.

Answer: .5443

e) The probability of 14 or less non-defectives.

Answer: .3828

f) The probability of 14 or more non-defectives.

Answer: .7858

g) The expected number of defectives.

Answer: 5

4. A contract has been negotiated between a consumer and a manufacturer for 1,000 spare parts at $5 each.

a) If the consumer considers it a serious risk to accept any lot in which 50 or more are defective and the manufacturer considers it a serious risk to reject any lot in which 30 or less are defective, what rejection number in a sampling plan of 100 should they agree on in order to balance their risks?

Solution: $n = 100$, $p_1 = .03$ $p_2 = .05$, $\alpha = \beta = ?$

		α	$1-\beta$
rejection c =	3,	.5802	.1183
	4,	.3528	.2578
	5,	.1821	.4360

Their risks are most closely balanced at rejection c = 4.

b) If, instead of balancing the risks, the consumer agrees that he would be willing to spend an additional $400 but no more on replacing defective parts previously accepted, what sampling plan should be used assuming that he wants to limit the risk of this $400 additional expenditure to .10?

Solution: $n = 100$, $p = \frac{400}{5000} = .08$, $\beta = .10$

Rejection $c = 5$, $\frac{1-\beta}{1-.9097} = .0903$

6, $1-.8201 = .1799$

The rejection number should therefore be $c = 5$

CHAPTER 11

Evaluation of Statistical Decision Rules in Terms of Expected Loss

11.1 The Terminal decision without sampling.[1]

Suppose a manufacturer is engaged in the manufacture of electron tubes for the U. S. Navy. Type 18 C 355 must be built to very close tolerances to have them perform satisfactorily. They require a coating of radioactive cobalt, grade A or grade B. Grade A cobalt is considerably more expensive than grade B. The tubes manufactured with grade A cobalt have 80 defectives per batch of 1000 tubes. Tubes manufactured with grade B cobalt are much more unpredictable. Previous experience has shown that the run characteristic using grade B cobalt has produced the following data which is considered valid.

Table 11-1

P	P(p)	No. of defectives per 1000
.12 defective	.12	120
.20 defective	.20	200
.35 defective	.35	350
.45 defective	.30	450
.55 defective	.03	550
	1.00	

[1] Bierman, Fouraker, and Jaedicke. Quantitative Analysis for Business Decisions. Pp. 161-171.

The tubes cost $80 apiece to manufacture using grade A cobalt. If grade B cobalt is used, the tubes cost $50 apiece. The additional manufacturing cost is caused by the difference in cost of the two grades of radioactive cobalt. Because of the rapid deterioration of the catalytic fusion agent, the radioactive cobalt is prepared in batches just large enough to fulfill the needs of each 1000 tube run. The cost of reworking a defective tube is $100.

If tubes come out on a run having a run characteristic of .12, .20, or .35, the manufacturer would want to use grade B radioactive cobalt because the overall costs are less than if he used grade A cobalt. If the tubes came out on a run having a run characteristic of .45 or .55, the manufacturer would want to use grade A cobalt because the overall costs are less than if he used grade B cobalt.

11.2 Computation of the breakeven fraction

a) Grade B cobalt cost per run:

Cost to rework = 1000 p x $100 = $100,000 p (1)

b) Grade A cobalt cost per run:

Cost to rework = 1000 x .08 x $100 = $8,000 (2)

Additional cost of Grade A = 1000 x $30 = 30,000

Total $38,000

The point at which the manufacturer would be indifferent as to which grade of cobalt to use is found by equating (1) and (2) above:

Table 11-2*

Expected costs of immediate Terminal Decision

		ACT			
		Use Grade A		Use Grade B	
p	P(p)	Cond. Cost	Expected Cost	Cond. Cost	Expected Cost
.12	.08	$38,000	$3,040	$12,000	$ 960
.20	.15	38,000	5,700	20,000	3,000
.35	.27	38,000	10,260	35,000	9,450
.45	.35	38,000	13,300	45,000	15,750
.55	.15	38,000	5,700	55,000	8,250
	1.00		$38,000		$37,410

*Note: The fixed costs of $50,000 for either act is omitted for simplicity.

$$\$100{,}000\ p_b = \$38{,}000$$

$$p_b = .38$$

(Note: p in reality could be other values than those listed in Table 11-1, however, in this example, they are not considered possible)

The manufacturer now is faced with a decision of whether he should sample or not sample the tubes coming off the production line on the first part of each 1000 tube run. Secondly, if he does decide to sample, what decision rule should he use? If the decision is not to sample, then Table 11-2 summarizes the Expected Cost of the two acts, i.e., 1) use grade A cobalt or, 2) use grade B cobalt. One can see that to make a terminal decision at this point he should use grade B cobalt because the expected cost is $590 cheaper than to use grade A cobalt. So a decision must be whether to act now on the basis of this information or to delay action in order to gather further information by sampling. The manufacturer, for example, could start using grade B cobalt and take a sample of tubes from the production line. Based on the number of defectives found in the sample, he would decide whether to continue this run with grade B cobalt or shift to grade A cobalt. Table 11-3 computes the Expected Cost under Certainty:

Table 11-3

Expected Cost under Certainty

P	P(p)	Best Cond. Cost Given each event	BEST ACT	Expected Cost
.12	.08	$12,000	Grade B	$ 960
.20	.15	20,000	Grade B	3,000
.35	.27	35,000	Grade B	9,450
.45	.35	38,000	Grade A	13,300
.55	.15	38,000	Grade A	5,700
	1.00			$32,410

The cost of uncertainty is $37,410- $32,410 = $5,000 and this sets the Expected Value of Perfect Information to the manufacturer.

If the cost of sampling is $100 per tube, it follows that the economic limit of n, size of the sample, is $5,000/100 = 50. Taking a sample larger than 50 would cost more than taking the lots as they came and using the best act, assuming that he somehow were clairvoyant enough to determine the characteristic of each lot before processing.

The next thing that must be done is to choose the best decision rule, i.e., the best n (sample size) and the best c (criterion number).

Table 11-4

Conditional costs

	ACT	
P	Use GRADE A	Use GRADE B
.12	$38,000	$12,000*
.20	38,000	20,000*
.35	38,000	35,000*
.45	38,000*	45,000
.55	38,000*	55,000

The conditional cost of each act is set forth in Table 11-4. The opportunity losses for these acts are set forth in Table 11-5. If the manufacturer wants to shift to Grade A cobalt when the process characteristic is .12, it will cost him $26,000 more. If he keeps using Grade B when the process characteristic is .55, it will cost him $17,000 more than had he shifted to Grade A cobalt, etc.

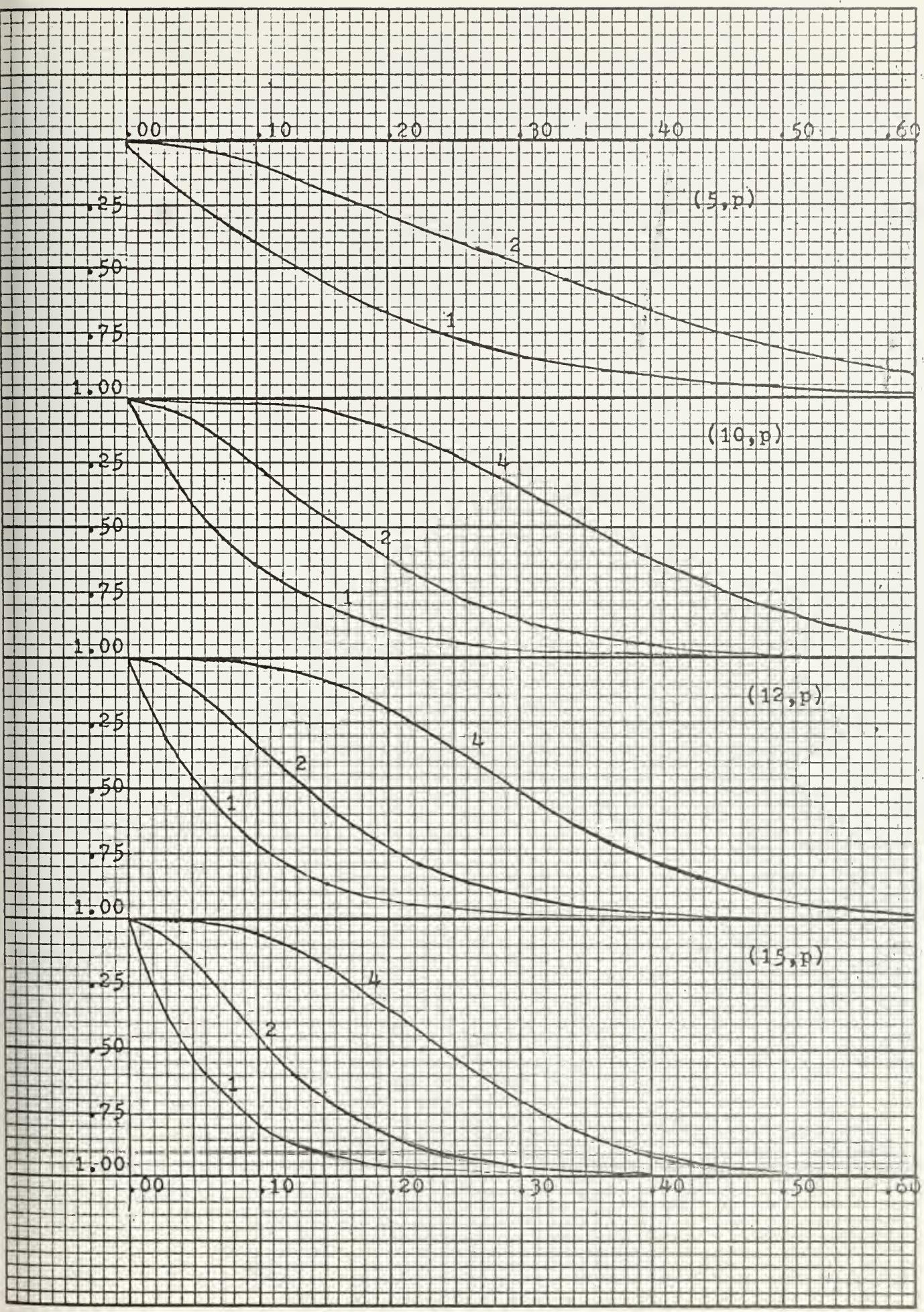
Table 11-5

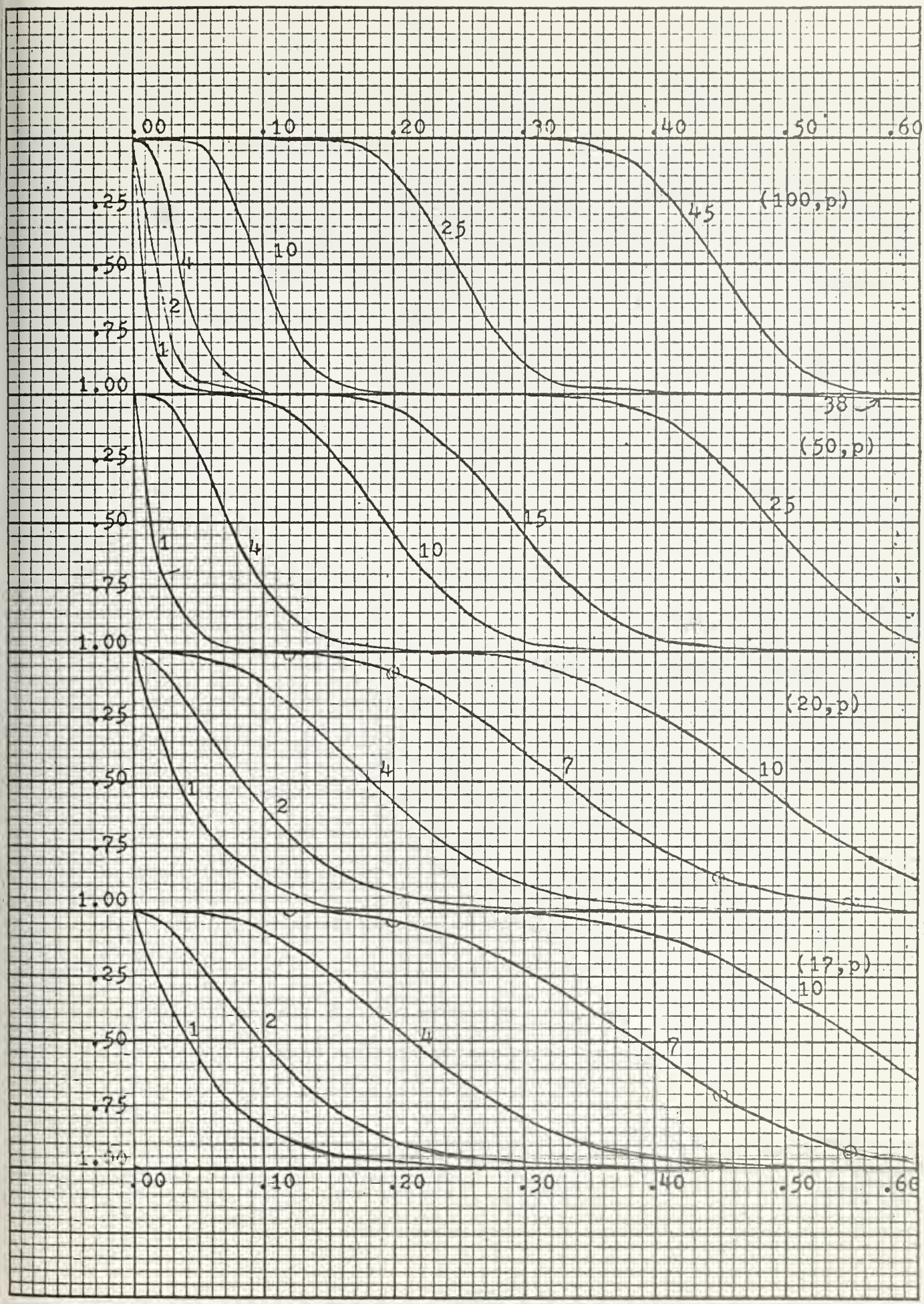
Conditional Opportunity Losses

	GRADE A	GRADE B			Column 6
.12	$26,000	$ 0	x .08	=	$2,080.00
.20	18,000	0	x .15	=	2,700.00
.35	3,000	0	x .27	=	810.00
.45	0	7,000	x .35	=	2,450.00
.55	0	17,000	x .15	=	2,550.00

Table 11-5 sets forth the opportunity losses if the wrong act is elected for fractions above and below the breakeven value, $P_b = .38$.

Figures 11-1 and 11-2 are selected error characteristic curves plotted for sample sizes 5, 10, 12, 15, 17, 20, 50, and 100 with various rejection numbers, and a p from 0 to .60. These curves are plotted directly from Table 1 in Schlaifer. It is possible to read the probability of wrong decision directly from these curves. To read the probabilities of wrong decision if the true p is less than the breakeven value, read the probabilities from the left-hand margin of the figure; if the true p is greater than the breakeven value then read the probability of wrong decision from the right-hand margin. For example, in the electron tube problem, for sample size 20, the probability of a wrong decision using a rejection number of 10 for a true p of .35 is P(wd) = .12. (It is read from the left-hand margin because .35 is less than $P_b = .38$). What we are saying is, that in a sample size of 20 tubes, if 10 or more are defective, the manufacturer in using the (20,10) decision rule, has a probability of .12 of making the wrong decision when p = .35, i.e., based on the sample he assumes that p of the batch is greater than .38 when it actually is .35, and rejects the null hypothesis $p \leq .38$, and instead of using Grade B cobalt he shifts to Grade A when he should have continued with Grade B; he would have made a TYPE I error (rejects the null hypothesis when true). To continue the illustration of decision rule (20,10), if the


.00
.10
.20
.30
.40
.50
.60
1.00
.75
.50
.25
(5,p)
2
1
(10,p)
4
2
1
(12,p)
4
2
1
(15,p)
4
2
1


.00
.10
.20
.30
.40
.50
.60
.25
.50
.75
1.00
1
2
4
10
25
45
(100,p)
38
(50,p)
15
(20,p)
7
(17,p)

true p is .45, the P(wd) = .60. It is read from the right-hand margin because p = .45 is greater than P_b = .38. In this instance, we are saying that in a sample of 20 tubes, if less than 10 are defective, the manufacturer in using (20,10) decision rule, has a probability of .60 of making the wrong decision when p = .45, i.e., based on the sample he accepts the null hypothesis that p of the batch is less than .38 when it actually is .45 and continues using grade B cobalt when he should have shifted to Grade A; he would have made a TYPE II error (accepts the null hypothesis when false).

11.3 Computation of optimal sample size

Decision rule (20,7).

The Total Expected Opportunity Loss will be the sum of the Expected Opportunity Losses for each fraction defective. The EOL is found by multiplying the Conditional Opportunity Losses by the joint probability of P(p) and P(wd). To give us more of a clue to which decision rule to start with, we multiply the column 6 by their respective P(p), as illustrated in Table 11-5, column 6. To obtain the smallest EOL, the P(wd) must be as small as possible for all fractions, but particularly for p = .12, .20, .45, and .55. Keeping in mind that the sample size must be less than 55, we can peruse Figures 11-1 and 11-2 for a reasonable decision rule and work from there. The rule (20,7) has been selected as a starting point since it seems to best fit the conditions stated above. The calculations are illustrated in Table 11-6.

Table 11-6

Decision rule (20,7) (Slide rule calculations)

p	Column 6		P(wd)		EOL
.12	$2,080	x	.0067	=	$ 13.95
.20	2,700	x	.0867	=	234.00
.35	810	x	.5834	=	472.00
.45	2,450	x	.1299	=	318.00
.55	2,550	x	.0214	=	54.60
					$1,092.55
			Cost of Sampling		2,000.00
			UETL	=	$3,092.55

UETL = Unconditional Expected Total (opportunity) Loss

Decision Rule (19,7)

p	Column 6		P(wd)		
.12	$2,080	x	.0[illegible]	=	$ 10.00
.20	2,700	x	.0676	=	182.30
.35	810	x	.5188	=	420.00
.45	2,450	x	.1727	=	423.00
.55	2,550	x	.0342	=	87.30
					$1,122.60
			Cost of Sampling		1,900.00
			UETL	=	$3,022.60

A tabulation of decision rules are given in Table 11-7.

Table 11-7

20,7	$3,092.55
20,8	3,179.42
19,7	3,022.60
19,8	3,233.68
19,6	3,110.80
18,7	3,000.58*
18,6	[illegible]
17,7	3,040.78

Based on the calculations which derived Table 11-7, it is concluded that the best decision rule is (18,7), i.e., n = 18 and c = 7. The opportunity loss theoretically will be reduced by $5,000 - $3,000), or $2,000 per run of electron tubes.

Note that in Table 11-5, column 6 the sum of value of column 6 for p = .12, .20, and .35 is the unconditional expected loss for the Terminal Act, Use grade A cobalt, and the sum of the values of column 6 for p = .45 and .55 is the unconditional expected loss for the Terminal Act, use grade B cobalt.

Problems

11.1[2] A manufacturer is faced with a problem of deciding how to inspect a particular part as it comes off the assembly line. The parts can either be good or defective, and the manufacturer can choose between two inspection systems. Both inspection systems will give perfect performance on good parts; that is, a good part will never be classified as defective. The two systems differ on classifying defectives. The first device, which is mechanical, will misclassify at a rate of .01, .05, 10, or .15, depending on how it is adjusted at the beginning of each production run. Because of the nature of the production run, once the machine is adjusted it cannot be readjusted until the next run. The adjustment record for the past 1,000 adjustments is as follows:

% Defectives Misclassified	Number of Occurrences
.01	100
.05	300
.10	500
.15	100
	1,000

The alternative inspection system is to place a workman on the assembly line, and the experience in the past has

[2] Ibid. Problem 12-4, p. 177. This problem requires a relatively large amount of time.

been that this system produces only .01 misclassifications.

If the workman is used, the incremental cost over the machine will be $25 per production run (a run is 1,000 pieces). However, when a defective is classified as good, there is a cost of reworking the subassembly when the part is used. This cost is $.50 per part misclassified.

a) Under what conditions would the inspection process be a Bernoulli process?

b) Assuming the conditions in (a) are fulfilled, calculate:

1) The expected monetary value of the best act.

2) The expected cost under certainty.

3) The expected value of perfect information.

c) Suppose the machine could be adjusted, a sample of up to 50 parts taken to check the adjustment, and then the machine could continue to be used or a workman could be placed on the assembly line for the remainder of the run.

1) Would a sample of 10 parts be worthwhile if the cost of sampling is $.10 per part? What is the best criterion number for a sample of 10?

2) Is a sample of 20 better than a sample of 10?

3) What is the maximum sample that would be taken, based on expected value of perfect information.

1. What are the probabilities that a destroyer will draw exactly 1, exactly 2 and exactly 3 reparables under the condition that there are

   a) 10 non-reparables in the lot?

   b) 35 non-reparables in the lot?

   c) 68 non-reparables in the lot?

2. What are the probabilities that a DLG will draw 9 or less, 6 or more, and less than 3 reparables under conditions

   1. (a), (b), and (c) above?

3. What are the probabilities that a CG will draw exactly 15 good, 8 or more, and less than three reparables under conditions

   1. (a), (b), and (c) above?

CHAPTER 12

Revision of Probabilities in the Light of New Information

12.1 The basic text gives an easy to follow explanation of Baye's theorem with figures and tables. Another approach is to view the problem in the same manner as that employed in Chapter 8 on marginal probability. Using the same illustration of section 12.1 in which we have an urn containing 10 balls in the following mix:

1. $\frac{1}{2}$ are striped; of these .2 are red and .8 are green
2. $\frac{1}{2}$ are dotted; of these .6 are red and .4 are green, we know the following:
   a) 10 balls, 5 are striped and 5 dotted
   b) of the striped ones, 1,(.2 x 5 = 1) is red and 4,(.8 x 5 = 4) are green
   c) of the dotted ones 3,(.6 x 5 = 3) are red and 2,(.4 x 5 = 2) are green

This can be set in a table of absolute values:

Table 12.1

	Red	Green	Total
Striped	1	4	5
Dotted	3	2	5
Total	4	6	10 10

and easily converted to a table of marginal probability:

Table 12.2

	Red	Green	Total
Striped	.1	.4	.5
Dotted	.3	.2	.5
Total	.4	.6	1.0 / 1.0

The four values in the center of the table are the various joint probabilities-red striped, red dotted, green striped and green dotted-while those on the periphery of the table are the unconditional probabilities-red, green, striped, and dotted.

The probability of drawing any one ball without prior information is one of the joint probabilities. But if a ball is drawn and it is known to be red, then the probability that it is striped is a conditional probability and is .1/.4 = .25. Given that a ball is dotted, the probability that it is green would be .2/.5 = .4. Etc.

The joint probabilities of the above table can be equated to the original probabilities of Figures 12.1 and 12.2 of the basic text. The revised probabilities of Figure 12.3 and Table 12.2 are obtained in a manner similar to that in the above paragraph.

The above illustration has practical applications in marketing and quality control because it reduces the cost of uncertainty which depends on the original prior prob-

abilities.

12.2 Assume that a manufacturer is turning out a machined product from raw material and that, dependent upon the quality of raw material, he has historically gotten a fraction defective as follows:[1]

Table 12.3

Fraction defective, p	Frequency of Occurrence P(p)
.01	.25
.05	.25
.10	.50
	1.00

It has previously been determined by the method shown in Chapter 11 that the best sample plan is (16,1) from a production run of 500 pieces. A sample is taken and two defectives found. How should the original prior probability distribution be revised and what action should the manufacturer take?

Two defectives in a sample of 16 is a conditional probability in that it depends on the fraction defective, p. Using the binomial table to compute the donditional probability of getting two defectives in a sample of 16 we get the following results:

[1]This example is derived largely from Bierman, Harold, Jr., Fouraker, Lawrence E., and Jaedicke, Robert K. Quantitative Analysis for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc. 1961. Pp. 161-175

Table 12.4

Fraction defective, p	Prior marginal probabilities known before sample was taken, P(p)	Conditional Probabilities from Sample $P(\tilde{r} = 2 \mid 16, p)$
.01	.25	.0104
.05	.25	.1463
.10	.50	.2745
	1.00	

We can now compute the joint probability for each fraction defective, i. e., the probability of a .01 defective batch of raw material and finding two defectives in a sample of 16, etc.

Table 12.5

Fraction Defective, p	Prior Prob. P(p)	Cond. Prob. $P(\tilde{r} = 2 \mid 16, p)$	Joint Prob. $P(\tilde{r} = 2, p)$
.01	.25	.0104	.0026
.05	.25	.1463	.0366
.10	.50	.2745	.1373
	1.00		.1765

The sum .1765 is the marginal probability and is the probability of getting two defectives in a sample of 16 from this raw material.

As a result of the sample information we would want to revise the prior probabilities. What we are looking for is the probability of the various fraction defectives given a

sample of 16 in which we found two defectives. This is a conditional probability and can be found in the same manner as above when we wanted the probability that a ball was striped when we had prior information that it was red; in that case we divided the joint probability .1 by the marginal probability .4 to get the conditional probability .25. Using the same method we would get the following revised (conditional) probabilities in the quality control problem:

Table 12.6

Fraction defective, p	Prior Prob. P(p)	Cond. Prob. $P(\tilde{r}=2 \mid 16,p)$	Joint Prob. $P(\tilde{r}=2,p)$	Rev. Prob. $P(p \mid \tilde{r}=2)$
.01	.25	.0104	.0026	$\frac{.0026}{.1765} = .0147$
.05	.25	.1463	.0366	$\frac{.0366}{.1765} = .2074$
.10	.50	.2745	.1373	$\frac{.1373}{.1765} = .7780$
	1.00		.1765	1.0001

Prior probabilities refer to the probabilities assigned to a basic random variable before a sample is taken and posterior probabilities refer to the revised probabilities as a result of sample information.

Problems[1]

1. A manufacturer assigns the following prior probabilities to the percent defective in a production process.

Event	Prob. of Event
.01 defective	.20
.05 "	.30
.10 "	.50

The rework cost of a defective is 50 cents and the length of the run is 1,000 pieces.

a) Compute the expected cost of using the process for a run of 1,000 pieces.

Solution:

Fraction Defective, p	Prob. P(p)	Cost: Conditional	Cost: Expected
.01	.20	$.50 x .01 x 1,000 = $5.00	$5.00 x .2 = $1.00
.05	.30	.50 x .05 x 1,000 = 25.00	25.00 x .3 = 7.50
.10	.50	.50 x .10 x 1,000 = 50.00	50.00 x .5 = 25.00
	1.00		$33.50

b) A sample of 10 pieces is run off and two defectives are found. Revise the probabilities to take account of the new information. Compute the expected cost of using the process.

[1]Problems 1 and 2 are derived from Bierman, Harold, Jr., Fouraker, Lawrence E., and Jaedicke, Robert K. Quantitative Analysis for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc., 1961 Pp. 178-179; problem 3 from Mosteller, Frederich, Rourke, Robert E. K., and Thomas, George B., Jr. Probability with Statistical Applications. Reading, Mass: Addison-Wesley Publishing Co., Inc., 1961 P. 148 and problem 4 from Wadsworth, George P. and Bryan, Joseph G. Introduction to Probability and Random Variables. New York: McGraw-Hill, 1960. P. 32

Problem 1 solution (cont'd)

Fraction Defective, p	Prior Prob P(p)	Cond. Prob $P(\tilde{r}=2 \mid 10,p)$	Joint Prob $P(\tilde{r}=2,p)$	Revised Prob $P(p \mid \tilde{r}=2)$	Expected Cost
.01	.2	.0042	.00084	.0072	.0072 x $5.00 = $ .036
.05	.3	.0646	.01938	.1656	.1656 x 25.00 = 4.140
.10	.5	.1937	.09685	.8273	.8273 x 50.00 = 41.365
	1.0		.11707	1.0001	$45.541

2. A manufacturer produces a special type of gear casting. His regular machine is in need of repair, but he has a standby machine which in the past has had the following defect record:

% defective	Historical frequencies
.01	100
.05	100
.10	800

For the current run of 500 parts the manufacturer can either subcontract the work at a cost of $10.09 per part or use the standby machine at an incremental cost of production of $10.00. The cost to rework a defective piece is $5.00.

a) What is the best act without sampling if the historical frequencies are considered prior probabilities?

What is the opportunity loss of each act?

Solution:

Fraction Defective, p	Prob P(p)	Cost due defectives Conditional	Expected
.01	.1	$5.00	$ .50
.05	.1	25.00	2.50
.10	.8	50.00	40.00
	1.0		$43.00

Expected cost of manufacture:

$$500 \times \$10.00 + \$43.00 = \$5,043.00$$

Cost of subcontract:

$$500 \times 10.09 = \$5,045.00$$

Therefore best act is to manufacture.

	Conditional Costs		Opportunity Losses	
Event	Manufacture	Sub-contract	Manufacture	Sub-contract
.01 defective	$5,005	$5,045	0	$40
.05 defective	5,025	5,045	0	20
.10 defective	5,050	5,045	$5	0

b) What is the best rejection number for a sample size 20?

Problem 2. solution (cont'd)

			(20,1)		(20,2)		(20,3)	
Event	Prob	Cond. Loss	Prob of Wrong dec.	Uncond. E.T.L.	Prob of Wrong dec.	Uncond. E.T.L.	Prob of Wrong dec.	Uncond. E.T.L.
.01 defective	.25	$40	.1821	$1.821	.0169	$.169	.0010	$ .010
.05 defective	.25	40	.6415	6.415	.0599	.599	.0755	.755
.10 defective	.50	5	.1216	.304	.3917	.979	.6769	1.692
				$8.540		$1.747		$ 2.457

Therefore the best rejection number is 2 for a Sample size 20.

3. In a factory machine A produces 30% of the output, machine B produces 25%, and machine C produces the rest. One percent of the output of machine A is defective, 1.2% of B's and 2% of C's. In a day's run the three machines produce 10,000 items. What is the probability that a defective item drawn at random from a day's output was produced by each machine?

   Answer: A- .2, B- .2, C- .6

4. An insurance company found that depositors who have sufficient funds in their checking accounts postdate a check by mistake once in a thousand times, while depositors who have insufficient funds invariably postdate their checks. When a postdated check is received at the bank, what is the probability that it was written by a depositor with insufficient funds if these depositors are one percent of the total?

   Answer: .91

CHAPTER 13

Suspension of Judgment and Summarization of Information

13.1 Analysis by Test of Significance[1]

Referring to the electron tube problem of Chapter 11, suppose a sample of 18 tubes has been taken and $r_o = 8$. Is the evidence derived from this sample sufficient to warrant a definite conclusion on the part of the manufacturer or should he take another sample before deciding that $p < .38$ and use grade B cobalt or that $p > .38$ and use grade A cobalt. By taking the breakeven value and multiplying by the sample size, $n = 18$, $p_b n = .38 \times 18$

$= 6.84$, but r must be discrete

so we use $r = 7$. Substituting in $P(\tilde{r} \leq p_b n - d_o | n, p_b)$ $+ P(\tilde{r} \geq p_b n + d_o | n, p_b)$ it follows that

$$P(\tilde{r} \leq 6 | n = 18, p = .38) = .4424$$
$$P(\tilde{r} \geq 8 | n = 18, p = .38) = .3681$$
$$\overline{.8105}$$

Similar to the example in 13.2.1 in Schlaifer, α would have to have been set at .8105, so the decision should be to take another sample. The probability of getting 6 or less or 8 or more defectives in a sample size of 18 at the breakeven value of p is .8105. This is also the probability of not getting 7 defectives.

[1]Robert Schlaifer, Statistics for Business Decisions. McGraw-Hill Book Co., Inc. 1961. p. 199.

13.2 Revision of prior probabilities given ($r_o = 8$, $n = 18$, p)

Having obtained 8 defects in our first sample, we can revise our prior probabilities with the information obtained from this sample and go through another analysis similar to that described in Chapter 11, to see if it is more economical to shift to grade A cobalt now or take another sample based on the revised probabilities. Table 13-1 shows the computation of the revised probabilities. Table 13-2 shows the computation for decision rule 20,7. Table 13-3 shows the computation for decision rule 19,7 and 20,8. Comparing the UETL from these three decision rules, it is readily seen that decision rule (20,7) is the best rule of the three, however we still don't know if it is the best rule for another sample since we have not tested decision rule (21,7). Here is one of the limitations to an incomplete set of binomial tables such as Schlaifer's Table I.

Table 13-1

Computation of Revised Probabilities

based on ($n = 18$, $\tilde{p} = p$, $r_o = 8$)

p	$P(p)$	x	$P(\tilde{r} = 8 \mid 18,p)$	=	$P(\tilde{r} = 8,p)$	$P(p \mid r = 8)$
.12	.08	x	.0005	=	.00004	.0003
.20	.15	x	.0120	=	.0018	.0147
.35	.27	x	.1326	=	.0358	.2940
.45	.35	x	.1863	=	.0653	.5370
.55	.15	x	.1247	=	.0187	.1540
					.12164	1.0000

Table 13-2

Decision rule (20,7)

P	$P(P\|\tilde{r} = 8)$		Cond. Loss		Column 4	x P(wd) =	EOL
.12	.0003	x	$26,000	=	$ 7.80	.0067 =	$ .05
.20	.0147	x	18,000	=	264.60	.0867 =	22.95
.35	.2940	x	3,000		882.00	.5834 =	514.00
.45	.5370	x	7,000		3,759.00	.1299 =	488.00
.55	.1540	x	17,000		2,618.00	.0214 =	56.00
							$1,081.00
					Cost of Sampling		2,000.00
					TOTAL		$3,081.00

Table 13-3

Decision rule 19,7

P	Column 4		P(wd)		
.12	$ 7.80	x	.0048	=	$ .04
.20	264.60	x	.0676	=	17.90
.35	882.00	x	.5188	=	457.00
.45	3,759.00	x	.1727	=	649.00
.55	2,618.00	x	.0342	=	89.50
					$1,213.44
			Cost of sampling		1,900.00
			UETL		$3,113.44

Decision rule 20,8

P	Column 4		P(wd)		
.12	$ 7.80	x	.0004	=	$ .03
.20	264.60	x	.0321	=	8.50
.35	882.00	x	.3990	=	352.80
.45	3,759.00		.2520	=	947.00
.55	2,618.00		.0580	=	152.00
					$1,460.33
			Cost of sampling		2,000.00
			UETL		$3,460.33

In order to test decision rule (21,7) a more complete set of binomial tables must be used. If they are not immediately available, they can be calculated by converting values derived from Table I, Schlaifer.

13-3 Calculation of a Probability when the sample size is increased by 1.

It is possible to use the Table I in Schlaifer to compute the $P(\tilde{r} = r|n \pm j,p)$. For example from Table I, $P(\tilde{r} = 7|19, .38)$ is .1860. The question is how can we calculate $P(\tilde{r} = 7|20,.38)$. Proceed as follows:

$$P(\tilde{r} = 7|19,.38) = C_7^{19}\, p^7\, q^{12} \tag{1}$$

$$.1860 = \frac{19!}{7!\ 12!} \times (.38)\ (.62)^{12} \tag{2}$$

$$P(\tilde{r} = 7|20,.38) = \frac{20!}{7!\ 13!} \times (.38)^7(.62)^{13} \tag{3}$$

$$= \frac{20 \times 19!}{7!\ 13 \times 12!} (.38)^7(.62)^{12}(.62) \tag{4}$$

$$= \frac{20 \times .62}{13} \times \frac{19!}{7!\ 12!} \times (.38)^7(.62)^{12} \tag{5}$$

$$P(\tilde{r} = 7|20,.38) = \frac{20 \times .62}{13} (.1860) \tag{6}$$

$$= .1774 \tag{7}$$

From Table I we read the same .1774. Therefore, by choosing the nearest value in the table, other values not given in a set of binomial tables may be computed in this same manner. The farther away from the table value, the more complex is the calculation however. To get the

cumulative probability, for example, $P(\tilde{r} = 7|21, .12)$, one would have to calculate each $P(\tilde{r} = p, 1, 2, 3, 4, 5, 6|20,.38)$ then sum these values and subtract from 1.0000. This process can become very laborious. A complete set of binomial tables is far more convenient. Table 13-4 is an example of such a calculation for $P(\tilde{r} \geq 7|21,.38)$.

Table 13-4

Calculation of $P(\tilde{r} = 7|21,.12)$ from $P(\tilde{r} = 0, 1, 2, 3, 4, 5, 6|20,.38)$

r	$P(\tilde{r} = r\|20,.12)$		Conversion Factor	$P(\tilde{r} = r\|21,.12)$
0	.0776	x	$\frac{.88 \times 21}{21}$	.0683
1	.1844	x	$\frac{.88 \times 21}{20}$	.1702
2	.2640	x	$\frac{.88 \times 21}{19}$	.2270
3	.2242	x	$\frac{.88 \times 21}{18}$	.2310
4	.1300	x	$\frac{.88 \times 21}{17}$	.1413
5	.0567	x	$\frac{.88 \times 21}{16}$	.0656
6	.0134	x	$\frac{.88 \times 21}{15}$	.0165
			$P(\tilde{r} < 7\|21, .12)$	= .9499
then			$P(\tilde{r} \geq 7\|21, .12)$	= .0501

Table 13-5 lists the values of P(wd) for decision rule (21,7) calculated by the method just illustrated.

Table 13-5

Probability (wrong decision 21,7)

p	P(wd)
.12	.0501
.20	.2017
.35	.6417
.45	.0964
.55	.0130

Table 13-6 is the summary of UETL's calculated for four decision rules. The best decision rule based on a revision of prior probabilities is (20,7). To determine whether or not the manufacturer should take this second sample, we must compare the Unconditional Expected Terminal Loss from decision rule (20,7) with the Expected Value of Perfect Information based on the posterior probabilities computed in Table 13-1. Table 13-, portrays the Expected Cost under Certain, for the posterior probabilities.

Table 13-6

Summary of UETL's

Decision Rule	UETL
19,7	$3,113.44
20,7*	3,081.00*
20,8	3,460.33
21,7	3,114.79

*Best rule

Table 13-7

Expected Cost under Certainty

P	P(p)	Best Cond. Cost Given each event	BEST ACT	Expected Cost
.12	.0003	$12,000	Grade B	$ 3.60
.20	.0147	20,000	Grade B	294.00
.35	.2940	35,000	Grade B	10,290.00
.45	.5370	38,000	Grade A	20,406.00
.55	.1540	38,000	Grade A	5,852.00
				$36,845.60

The cost of uncertainty is

$38,000 - $36,846 = $1,154

and this sets the Expected Value of Perfect Information to the manufacturer. By comparing the EVPI with the UETL of decision rule (20,7), i.e.,

EVPI = $1,154

UETL = 3,081,

it is apparent that the manufacturer should not take the second sample, but switch to grade A cobalt and complete the run since the UETL exceeds the EVPI by $1,927.

Problem[1]

13-1 A manufacturer assigns the following prior probabilities to the per cent defective in a production process considered Bernoulli:

Event	Probability of Event
.01 defective	.20
.05 defective	.30
.10 defective	.50
	1.00

The rework cost of a defective is 50 cents; the length of run is 1,000 pieces.

a) Compute the expected cost of using the process for a run of 1,000 pieces.

b) A sample of ten pieces is run off and two defectives are found. Revise the probabilities to take account of the new information. Compute the expected cost of using the process.

[1]Harold Bierman, and Others, Quantitative Analysis for Business Decisions. p. 179, Problem 12-8.

CHAPTER 14

Measures of Variability; Variance and Standard Deviation

14.1 General

There is a common propensity for like items to locate themselves about a central point. In Chapter Six, we discussed briefly some measures of central tendency; namely, the mean, median and mode. But measures of *inequality* such as distribution of income, IQ of students at the U. S. Naval Postgraduate School, height by age groups of Japanese people, educational level of the Russian people, etc., are problems requiring the measurement of *variability.*

While the mean, median and mode describe respective central locations in a distribution, they tell us nothing about the variability of a distribution. For example, if we knew that the mean depth of a river was one foot, we had better not try wading across until we find out what the extreme depths are, for the mean does not really tell us much about the river depth. What we want to know is how representative is the mean; how is the distribution spread around the mean? What is the "spread", "scatter", "dispersion" or *variability?* The answer to this question gives rise to two new terms: the *variance* and its square root, the *standard deviation.*

14.2 The Variance

The difference between the mean and an individual value in a distribution is the deviation of that value.

The variance is found simply by (1) averaging the squares of all deviations in the distribution or (2) by summing the weighted (i.e., relative frequency) deviations squared.

Using the data of Table 6.1 Hypothetical Distribution of Lieutenants in the Navy, we know that the mean of the random variable, $\tilde{Z}$ is 27; we can compute the variance, $\sigma^2(\tilde{Z})$, in either of two ways as follows:

Table 14.1

			Ave. of deviations (1)		Wt'd. ave. method (2)	
Value of the random variable $\tilde{Z}$	Deviation	Deviation Squared	No. of occurrences	Product	Rel. Freq	Product
23	-4	16	750	12,000	.05	.80
24	-3	9	750	6,750	.05	.45
25	-2	4	1,500	6,000	.10	.40
26	-1	1	2,250	2,250	.15	.15
27	0	0	3,750	0	.25	0
28	+1	1	3,000	3,000	.20	.20
29	+2	4	1,500	6,000	.10	.40
30	+3	9	750	6,750	.05	.45
31	+4	16	750	12,000	.05	.80
			15,000	54,750	1.00	3.65

$$\text{Variance} = \sigma^2(\tilde{Z}) = \frac{54,750}{15,000} = 3.65$$

14.3 The Standard Deviation

The Standard Deviation is simply the square root of the variance.

In the above example, Standard Deviation

$$\sigma(\tilde{Z}) = 3.65 = 1.91 \text{ years}$$

The variance and standard deviation are measures of the variability of a distribution which have not previously been used in the basic text. Although they are not the only measures of variability, they are the only ones suitable for the sampling type problems found in the subsequent chapters of the text. Furthermore, we have to this point been concerned only with the binomial distribution which has a serious limitation in that it requires an unweildy number of tables in order to be used in all problems. This limitation can be met by use of the Normal approximation to binomial probabilities; complete Normal probability tables occupy only one page. The term "Normal" also extends to the distribution itself which has certain characteristics that are described in the next chapter. It also has some limitations; in general, in order to use the Normal distribution as an approximation to the binomial,

$$p \times n \geq 5$$

$$q \times n \geq 5$$

That is, the mean long run expected value of successes and failures in the Bernoulli process should be equal to or

greater than five.

Problems:

1. Find the approximate mean, variance and standard deviation by two methods of wages in the following group of 500 workers.

WAGES	NO. OF WORKERS
$ 40 to $ 60	60
60 " 80	70
80 " 100	85
100 " 120	110
120 " 140	95
140 " 160	80
	500

Answer: Mean $104, Variance $1,020, Standard Deviation $31.90.

2. One automatic screw machine is used to produce shafts and another to produce bushings one of which is assembled over each shaft. The mean diameter of the shafts is 1.000 inch; the mean inside diameter of the bushings is 1.002 inches; and the Standard deviation of either set of diameters is .001 inch. Compute the mean clearance (bushing diameter minus shaft diameter) and the standard deviation of the distribution of clearance on the assumption that assembly is random--i.e.,

bushings are not selected to fit the shafts on which they are placed.

Answer:

Mean clearance, .002; stand deviation, .001414[1]

[1]This problem is from Schlaifer, Robert. Probability and Statistics for Business Decisions. New York: McGraw Hill Book Company Inc., 1959. P. 272

CHAPTER 15

The Normal Distribution and the Normal Approximation of the Binomial Distribution

15-1 Approximating the Binomial Distribution

$P(r = r \mid n = 50, p = .5)$.

The normal distribution is an excellent model and useful addition to the statistician's tool kit. It can be used to approximate the binomial distribution. It should be remembered that the binomial distribution is an exact discrete distribution. The Normal distribution is a continuous distribution, so in order to approximate the binomial distribution, it must be made discrete by approximating a binomial histogram. For example, take the binomial distribution $P(\tilde{r} \geq r \mid n = 50, p = .5)$, how can we approximate this distribution with the Normal? Proceed as follows:

$n = 50$ $\qquad \sigma^2 = npq$

$p = .5$ $\qquad \sigma = \sqrt{npq}$

$q = .5$ $\qquad \sigma = \sqrt{50 \times .5 \times .5}$

$E(\tilde{r}) = pn$ $\qquad \sigma = 3.53$

$= .5 \times 50$

$E(\tilde{r}) = 25$

For the Normal distribution,

$$u = \frac{r - 25}{3.53} .$$

By calculating the values of u and plugging Table III in Schlaifer, the data in column 3, Table 15-1 was derived.

Column 2 is comparative data taken from Table I in Schlaifer. Notice how very close the Normal distribution approximates the binomial under the particular parameters of (n = 50, p = .5).

Two examples of the calculations are given below with Fig. 15-1 illustrating the logic.

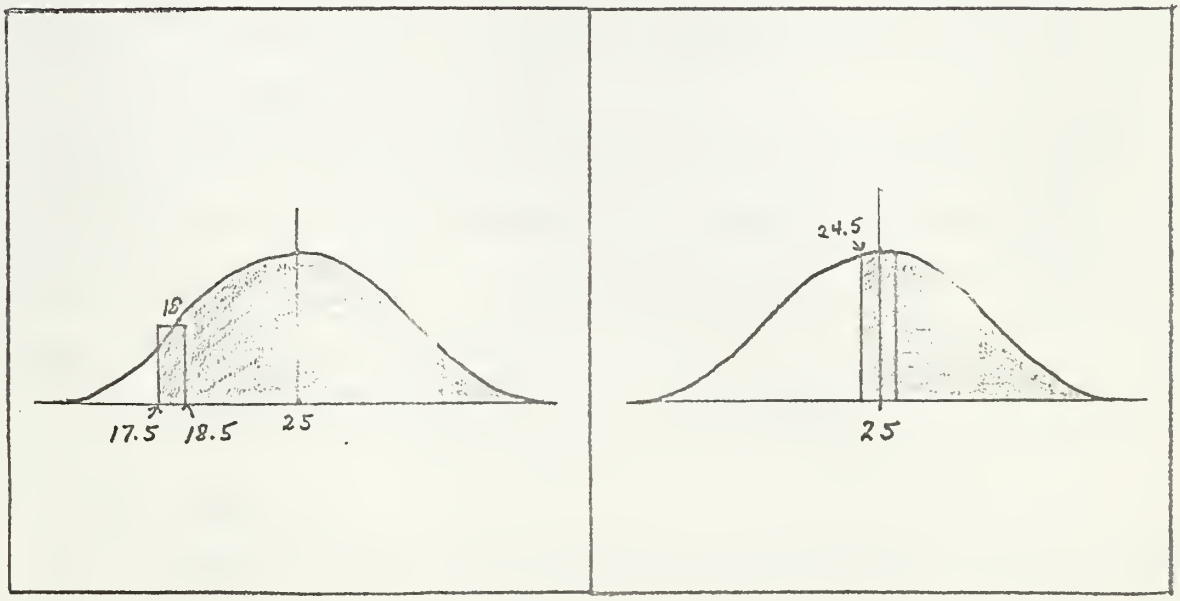
(1) For $P(\tilde{r} \geq 18 | 50, .5)$

$$u = \frac{17.5 - 25}{3.53}$$

$$u = -2.12$$

From Table III in Schlaifer, read .01700. Then 1.0000-.01700 = .9830.

Fig. 15-1



From Fig. 15-1, it can be seen that the shaded area is what is desired. The unshaded area is read directly from Table III. Since the total area under the curve is 1.0000, to obtain the shaded area, all that is necessary is to subtract the unshaded area from 1.0000. As Schlaifer has pointed out, the Normal curve is symmetrical with respect to the mean so that it is only necessary to provide tables for half the Normal distribution and the other half can readily be derived.

(2) For $P(\tilde{r} \geq 25 | 50, .5)$

$$u = \frac{24.5 - 25}{3.53}$$

$$u = -.1416$$

Enter Table III in Schlaifer read and interpolate .4439,

then $P(\tilde{r} \geq 25 | 50, .5) = (.5000) - .4439 + .5000$

$$P(\tilde{r} \geq 25 | 50, .5) = .5561$$

although Table 15-1 portrays $P(\tilde{r} \geq r)$ for r from 12 to 39, one should realize that there are binomial probabilities for $r = \emptyset, 1, 2, 3, \ldots 11$, but they are so small as to be almost trivial. For example

$$P(\tilde{r} = \emptyset | 50, .5) = \frac{50!}{0!\ 50!} \times \quad)^0 (.5)^{50}$$

$P(\tilde{r} = \emptyset | 50, .5)$

$= .000\ 000\ 000\ 000\ 001\ 776\ 356\ 839\ 400\ 250\ 464\ 677\ 810\ 668\ 945\ 312\ 5$

Table 15-1

r	Binomial $P(\tilde{r} \geq r \mid 50,.5)$	Normal $P(\tilde{r} \geq r \mid 50,.5)$	Difference
12	1.0000 - e *	.9999	- .0001
13	.9998	.9998	.0000
14	.9995	.9994	- .0001
15	.9987	.9985	- .0002
16	.9967	.9964	- .0003
17	.9923	.9919	- .0004
18	.9836	.9830	- .0006
19	.9675	.9671	- .0004
20	.9405	.9406	+ .0001
21	.8987	.8988	+ .0001
22	.8389	.8389	.0000
23	.7601	.7601	+ .0001
24	.6641	.6642	- .0001
25	.5561	.5561	.0000
26	.4439	.4439	.0000
27	.3359	.3359	.0001
28	.2399	.2399	.0000
29	.1611	.1611	.0000
30	.1013	.1012	- .0001
31	.0595	.0596	+ .0001
32	.0325	.0329	.0004
33	.0164	.0170	+ .0006
34	.0077	.0080	+ .0003
35	.0033	.0036	+ .0003
36	.0013	.0015	+ .0002
37	.0005	.0006	+ .0001
38	.0002	.0003	.0001
39	.0000	.0001	+ .0001

* a very small number.

when p = .5 the Binomial distribution is symmetrical with respect to $E(\tilde{r})$, and the Normal distribution approximates it even when n is fairly small for example n = 10. Table 15-2 illustrates this point.

Table 15-2

r	Binomical $P(\tilde{r} \geq r \mid 10,.5)$	Normal $P(\tilde{r} \geq r \mid 10,.5)$	Difference
0	1.0000	.9997	- .0003
1	.9990	.9991	+ .0001
2	.9893	.9864	- .0029
3	[illegible]	[illegible]29	- .0024
4	[illegible]	.[illegible]8	+ .0008
5	.623[illegible]	.624[illegible]	+ .0010
6	.37[illegible]	.37[illegible]	- .0010
7	.1719	.1711	- .0008
8	.0547	.0571	+ .0024
9	.0107	.0136	+ .0029
10	.0010	.0009	- .0001

When p or q are small the bino[illegible]istribution skews right or skews left and the Normal distribution approxima[illegible]on is not as good as the Poisson distribution which also can approximate the binomial as pointed out in Schlaifer. The graphs in figures 15.6, 15.7, and 15.8 in Schlaifer bear this out.

Problems:

15-1 Verify the data in Table 15-2.

15-2[1] A fair coin is tossed 500 times. Find the probability that the number of heads will not differ from 250 by (a) more than 10, (b) more than 50.

15-3[2] If 10 per cent of television picture tubes burn out before their guarantee has expired, (a) what is the probability that a merchant who has sold 100 such tubes will be forced to replace at least 20 of them? (b) What is the probability that he will replace at least 5 and not more than 15 tubes? Use the normal curve approximation.

---

[1]Murray R. Spiegel, Theory and Problems of Statistics. (Schaum Publishing Co., 1961) Prob. 25, p. 135.

[2]Paul G. Hoel, Elementary Statistics. (John Wiley & Sons Inc., 1960) Prob. 15. p. 69.

CHAPTER 16

The Central Limit Theorem and Large-Sample Theory

16.1 Background

In earlier parts of Schlaifer's text considerable attention was paid to principles of mathematical probability. Indeed, we have periodically returned to these principles throughout the course. One of the most important uses of mathematical probability in statistics is its application to sampling theory. Not only are sampling techniques used extensively in business problems as stressed in the basic text but also in social, biological and physical sciences, government, and humanities. However, statistics and sampling techniques have definite limitations-sample size in relation to the total population, representativeness, definitions, inaccurate measurement or classification, inappropriate comparisons, disregard of dispersion, etc.--when used or interpreted improperly. In fact, a book has been written on How To Lie with Statistics[1] and at least one text contains a whole chapter on "Misuses of Statistics".[2] Two books which became popular best-sellers in the 1950's were

[1]Huff, Darrell. How To Lie with Statistics. New York: W. W. Norton & Company, Inc., 1954.

[2]Wallis, W. Allen and Roberts, Harry V. Statistics a New Approach. The Free Press of Glencoe Inc., 1962, Chapter 3.

the Kinsey Reports.[1] But both of these volumes have been roundly discredited by statisticians who cut through the sensational findings to examine how they were obtained. The sampling techniques were so inadequate to most statisticians that confidence in the basic significance of the findings was diminished such that the books have little value beyond their role in opening a broad and important field.[2]

Sampling, therefore, is an important part of statistics. How do we go about obtaining samples, how representative are they, how much freedom are we allowed in interepretation and mathematical manipulation, what inferences can be drawn from samples, how confident are we that a sample represents the true population from which drawn, what is the significance of a sample, and what conclusions can be drawn?

16.2 Symbols

Listed below are some of the more important symbols used in sampling theory:

x - an individual value in a distribution (example: the wage, $2.00/hr. of one man in a factory)

---

[1]Kinsey, Alfred C., Pomeroy, Wardell B., and Martin, Clyde E. Sexual Behavior in the Human Male. Philadelphia and London: W. B. Saunders Company, 1948 and Kinsey, Alfred C. and Associates. Sexual Behavior in the Human Female. New York: E. P. Dutton and Co., 1954.

[2]Wallis, W. Allen. "The Statistics of the Kinsey Report." Journal of the American Statistical Association. Vol. 44 (1949) P. 466.

$\bar{x}$ - the mean of a sample (example: from a sample of, say, 5 men whose wages are \$1.80, \$1.95, \$2.00, \$2.20 and \$2.25, in a total of 100 in a factory a mean wage, \$2.04, is computed)

t - the total sum of x's (individual values) in a sample (example: the sum, \$10.20, of the wages in the sample of 5)

μ - the mean of all the x's (individual values) in a population (example: the mean of all the 100 wages in the factory)

σ - the standard deviation of all the x's (individual values) in a population

$\sigma_{\bar{x}}$ - the Standard deviation of the sample mean (example developed below)

N - size of a population, the number of x's in a population (example: N = 100 wages of workers in the above cited factory)

n - number of items in a sample (example: n = 5 above)

16.3 Sampling with and without replacement

If we have an urn with a total of 100 red, white, and blue balls in it and start drawing out a sample of 12 in order to estimate the true proportion of each, then the probabilities of drawing each color will change after every draw. The probability of drawing any one color is

conditional upon its true parameter, p, and the ones already drawn. This is called sampling without replacement and often is referred to as sampling from a finite population. If we replace a ball each time it is drawn for our sample of 12, then each draw has no bearing on the one that comes after it and the probability of drawing any one color is conditional upon its true population parameter only. While the size of the population may be finite, it can be considered infinite because any number of samples can be drawn with replacement without exhausting the population. There are many practical problems in which sampling is done from a finite population but is treated as infinite because of its large size.

16.4 The Central Limit Theorem

Although a distribution can have a graphical shape other than Normal, as illustrated in Figure 16.1 of the basic text, most of the problems we will be concerned with approach a Normal distribution. In this connection, however, there is a tendency for sampling distributions to assume a Normal distribution. This is ordinarily true regardless of the shape of the population from which the sample was drawn, and the larger the sample, the more closely does it approach Normality. This tendency is called the Central Limit Theorem and its importance lies in the fact that it allows us to apply the Normal distribution to

statistical sampling distributions.

Except in the comparitively unusual gamma and exponential processes, a sample will approximate a Normal distribution if: (1) the sample itself is large enough in absolute size to approximate a Normal distribution; that is, $n \geq 30$ (according to most texts) and (2) if the population size, N, is large enough (at least twice the sample size) to make the sample virtually independent of the population.

16.5 The Finite Population Correction

If the sample is not independent of the population; that is, if the sample takes in more than 5% of the population, then the finite population correction should be applied, using the formula as stated in Schlaifer:

$$\sigma(\bar{t}) = \sigma\sqrt{n}\sqrt{1-\frac{n}{N}}$$

or as stated in other texts:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{1-\frac{n}{N}}$$

If the sample is less than 5% of the population, which is the usual case, then (1) the size of the population will not affect the Normality of the sample providing the sample itself is large enough and (2) the finite-population correction can be disregarded in computing the standard deviation of the Normal approximation since it is neglible. Thus in most practical applications it is safe to assume that the distribution of sample means is Normal with a standard

deviation of $\bar{x} = \frac{\sigma}{\sqrt{n}}$.

16.6 Estimating the population Standard deviation and degrees of freedom

Thus far we have seen that three Statistical theorems give us virtually complete information about sampling distributions:

1) The central limit theorem states that a sample will be approximately Normal if large enough ($n \geq 30$),

2) The mean, $\bar{x}$, of the sample means is always equal to the population mean; that is,

$$\bar{x} = \frac{\sum(x)}{n} = \mu = \frac{\sum(x)}{N}$$, where the first sum is the values in the sample and the second sum is the values in the population,

3) The Standard deviation of the sample means is always a certain multiple of the standard deviation of the population; that is,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$.

Thus, both the mean and standard deviation of the sample depend on the corresponding values of the population. However, in most practical situations these values are not known. In fact, one reason for sampling is to make an estimate of what these values are in the population.

The above theorems hold true for large samples

($n \geq 30$), but not for small. This can be seen intuitively if you consider a sample of size one which obviously would not be equal to the population mean except by coincidence. A sample value is likely to be closer to its own mean (particularly when the sample is small) than to the population mean. This again is intuitively apparent, for a small sample will not likely reproduce the range of a population. Consequently the standard deviation of a small sample will tend to be lower than the standard deviation of the population. This factor can be compensated for by applying a correction to the sample in order to make it a more realistic estimate of the population. Since the standard deviation of the sample tends to be smaller than that of the population, we want to enlarge it; and this can be done by using a larger divisor than n in the formula for the standard deviation of a sample; thus for small samples the estimate of the population standard deviation,

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} .$$

The degrees of freedom are the number of items in a sample that are effective in estimating a specified characteristic of a population.[1] Using the factor (n-1) results in one degree of freedom being lost. Examining the above formula it can be seen (as previously stated) that decreasing the denominator from n to (n-1) will increase the standard

[1]Smith, Frank C. and Leabo, D. A. Basic Statistics for Business Economics. Homewood, Illinois: Richard D. Irwin, Inc., 1960, P. 94.

deviation when the sample is small but make little difference when it is large. For most practical purposes we can consider that the standard deviation of the population, σ, has the same value as its estimate, s, providing that the sample from which s is obtained has at least 25 degrees of freedom.

Problems[1]

1. A sample of ten college graduates earned a mean of $2,000 per year more than the average earnings of high school graduates with a standard deviation of $500 more. What are the 95 per cent confidence limits of the mean amount by which all college graduates' annual income exceeds the mean annual earnings of high school graduates?

   Answer: $1,623-$2,377.

2. A man claims he is a 170 bowler. For 49 games he has a mean score of 165 with a standard deviation of 21. If you reject his claim, what is the probability that you have made a mistake?

---

[1]Problems 1-4,8,9 and 10 are taken from Perlman, Richard. Problems in Statistics for Economics and Business Students. New York: Holt, Rinehart and Winston, Inc., 1963; problems 5-7 are from Mosteller, Frederick, Rourke, Robert E. K., and Thomas, George B. Jr. Probability with Statistical Applications. Reading, Mass: Addison-Wesley Publishing Company, Inc., 1961; and problem 11 from Neter, John, and Wasserman, William. Fundamental Statistics for Business and Economics. Boston: Allyn and Bacon, Inc., 1961.

Answer: .048

3. A company conducting surveys of consumer acceptance of new product pays its field force by the interview. It suspects that the interviewers, who average 20 interviews per day, are spending too little time on interviews. A company official goes in the field and over a period of 10 days averages 18 interviews per day with a standard deviation of five per day. Using the one per cent level of significance, should he conclude that the field force hurries its interviews?

   Answer: No

4. Golfer A knows he averages 80 strokes per round. He observes that Golfer B averages 86 strokes for 10 rounds with a standard deviation of 6 per round. How many strokes handicap per round can Golfer A give Golfer B and feel 95 per cent certain of beating him in the long run?

   Answer: 2.33 strokes

5. From a population size N = 40, a sample is to be drawn, and the sample average is to be used to estimate the population mean. Which provides a more accurate estimate, a sample of 8 drawn without replacement or a sample of 10 drawn with

replacement and why?

Answer: 10 with replacement because $\sigma^2/10$ is less than $4\sigma^2/39$

6. A certain mental test yields scores in months of mental age. The errors of measurement in this test average zero in the long run, with a standard deviation of 2 months. If a class of 36 students takes the test, what is the probability that the average score for the class is in error by one month or more?

   Answer: .0026

7. In crossing two pink flowers of a certain variety, the resulting flowers are either white, red, or pink with respective probabilities of $\frac{1}{4}$, $\frac{1}{4}$, and $\frac{1}{2}$. If 300 flowers are obtained by crossing pink flowers of this variety, what is the probability that 90 or more of these flowers are white?

   Answer: .0228

8. A drive aims for a goal of one million dollars from a college's 20,000 alumni. A sample of 26 responses yields an average contribution of $40 with a standard deviation of $20. Based on the sample information, is there at least a 5% probability of reaching the goal?

   Answer: No. The maximum average received is $46.83

and the average required is $50.00.

9. For a certain occupation, past experience indicates a standard deviation of wage of $.50 per hour. How large must a sample of workers who average $2.50 be in order that the 95% confidence interval has a range of $.25?
   Answer: 61

10. A research report stated that the mean income per person in the population was $4,500 ± $800, with a confidence level of 95%. A student interpreted this to mean that 95% of the persons in the population had incomes between $3,700 and $5,300. Another student interpreted this to mean that if many samples were taken, 95% of them would have sample means between $3,700 and $5,300. What basic mistakes did both students make?

CHAPTER 17

Statistical Decision Rules with Normal Sampling

17.1 General The problems discussed in Chapter Seventeen of the basic text are similar to those of Chapter Ten except that they are based on the Normal distribution instead of the binomial. The concept of the Null and Alternate Hypotheses, Type I and Type II Errors and Producer's and Consumer's Risks are equally applicable here.

17.2 Formulating a Decision Rule

The decision rule is determined in the same manner as previously, i.e., take a sample of size n (which must be sufficiently large for a Normal distribution) and compute the mean $\bar{x}$ of this sample. If $\bar{x}$ is greater than c accept the process. The key to this procedure is the selection of the decision rule or (n,c) pair. Just as we observed in Chapter Ten, the value of c can be changed while keeping n fixed and thereby decrease the probability of a Type I error, but only at the expense of increasing the probability of a Type II error or vice versa depending on the statement of Hypotheses. The probability of both type errors can be reduced by increasing n and changing c appropriately but this incurs increased sampling cost. However, as a practical matter it is so difficult to formulate a decision rule by comparing the conditional probabilities of wrong decision

for all values of μ under all possible (n,c) combinations that the decision rule is usually made by accepting certain specified levels at which one is willing to be in error. This procedure is outlined on page 276 of the text and is almost identical to that contained on page 158 where the procedure for deriving a decision rule under the Binomial Distribution was developed.

17.3 The Error Characteristic The error characteristic of a decision rule is the conditional probability, given any μ , that the decision rule will lead to a wrong decision, that is, that a Type I or II error will be made. Figures 17.2, 17.3, 17.4 and 17.5 of the text illustrate the error characteristics of several decision rules.

17.4 Setting the Conditional Probabilities of Error In setting the conditional probability of error that the decision maker is willing to accept, the greater his requirement for accuracy, and the less will be his acceptance of error. Thus a test that is significant at the .01 level has a greater probability of reflecting the true parameter of the population than a test at the .05 level. In any particular decision rule of this type the sum of the confidence level and the significance level is one. Sometimes it is desireable that the α and β risks be balanced in which case they are set at equal levels. On the other hand in some processes it may be considered that changing the status quo

for the worse is a more serious error than failing to change it for the better; in other words, altering a good process may be more serious than failing to alter a bad one. For this reason, there are many processes in which α and β are deliberately not balanced, and the α risk is deliberately set at a lower level than the β risk. Such is the case in an illustrative problem that follows.

17.5 Controlling alpha and beta on a two-tailed test[1]

A manufacturer is boxing corn flakes in family size packages. The contents should weigh 16 ounces. Management wishes to set up a scheme to control the process; they want neither too much nor too little cereal in the box. If too much is put in it reduces profits and if too little is put in they cheat the customer. What can the company statistician do to help set up the controls?

1) He can recommend a type of sampling design.

2) He can determine the proper sample size.

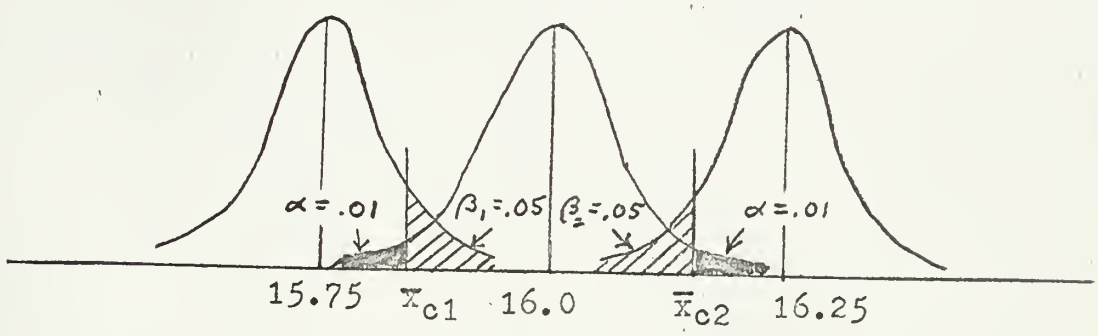
3) He can set up a decision rule.

Assuming that the mean μ of the process is 16 ounces and the standard deviation of the contents of the package is .4 ounce, management wishes to limit the probability of making a Type I error (alpha risk) to two per cent. Management also desires to limit the β_1 risk to 5 per cent of

[1]Frank C. Smith and D. A. Leabo. Basic Statistics for Business Economics. Pp. 156-161.

concluding that the boxes contain 16 ounces on the average when the true mean is 15.75 ounces and the β_2 risk to 5% of concluding that the boxes contain an average of 16 ounces when the true mean is 16.25 ounces. What size sample and what $\bar{x}_c$ s should the manufacturer employ to fulfill the given criteria. Figure 17-1 is a diagram which illustrates the problem.

Figure 17-1



17.6 Computation of Sample size and $\bar{x}_{c1}$ and $\bar{x}_{c2}$.

$$\frac{\bar{x}_{c1} - 16}{.4/\sqrt{n}} = -2.33 \;\Big\|\; \alpha = .01 \qquad (1)$$

$$\frac{x_{c1} - 15.75}{.4/\sqrt{n}} = 1.645 \;\Big\|\; \beta_1 = .05 \qquad (2)$$

Solving (1) and (2) simultaneously for $\bar{x}_{c1}$ and n we have,

from (1),

$$\bar{x}_{c1} - 16 = \frac{-.4 \times 2.33}{\sqrt{n}}$$

$$\sqrt{n} = \frac{-.4 \times 2.33}{\bar{x}_{c1} - 16}$$

from (2), substituting for $\sqrt{n}$,

$$\bar{x}_{c1} - 15.75 = \frac{.4 \times 1.645}{-.4 \times 2.33} (\bar{x}_{c1} - 16)$$

$$1.707\bar{x}_{c1} = 27.05$$

$$\bar{x}_{c1} = 15.85 ,$$

then

$$\sqrt{n} = \frac{-.4 \times 2.33}{15.85 - 16}$$

$$\sqrt{n} = \frac{-.932}{-.15}$$

$$\sqrt{n} = 6.22$$

$$n = 38.7$$

and since n must be an integer,

$$n = 39.$$

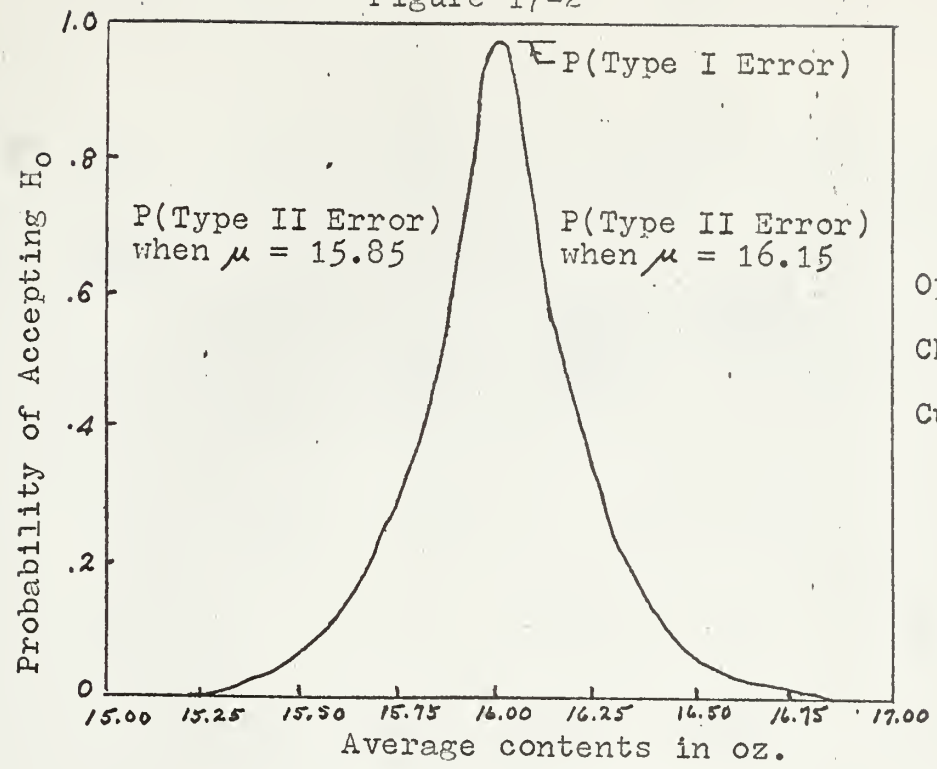
To compute $\bar{x}_{c2}$,

$$\frac{\bar{x}_{c2} - 16}{.4/6.22} = 2.33$$

$$\bar{x}_{c2} = 16.15.$$

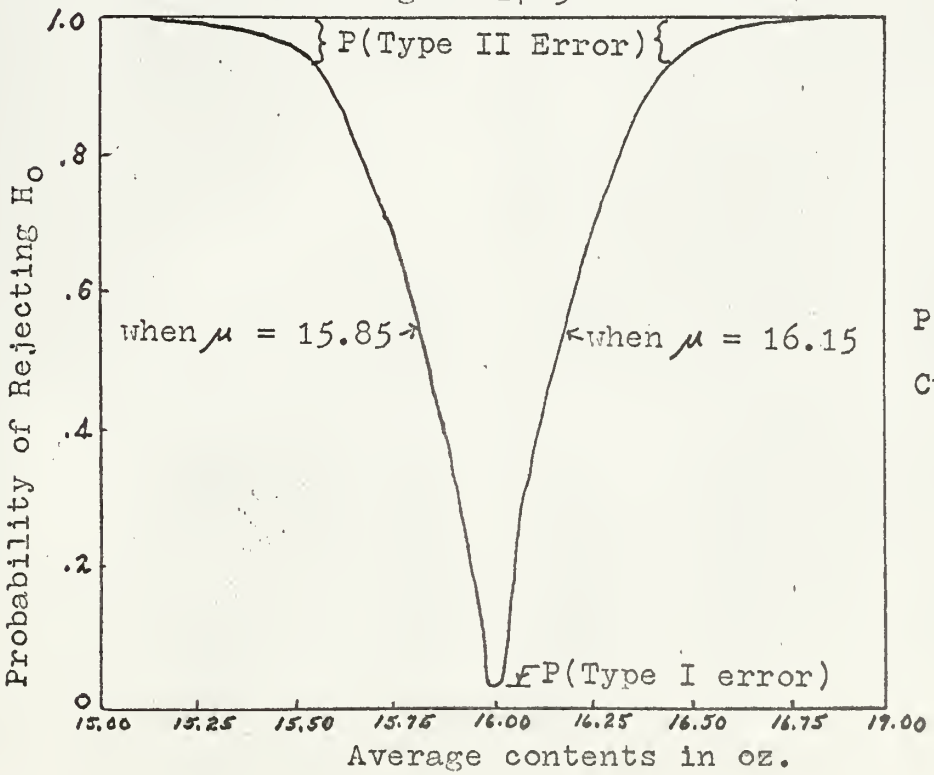
Consequently the sampling plan is to select a sample of size 39, if the mean contents of these 39 boxes, $\bar{x}$ is

Figure 17-2[2]



Operating Characteristic Curve

Figure 17-3[2]



Power Curve

[2]Ibid. Pp. 160-161.

Calculation of data in Table 17-1 is as follows:

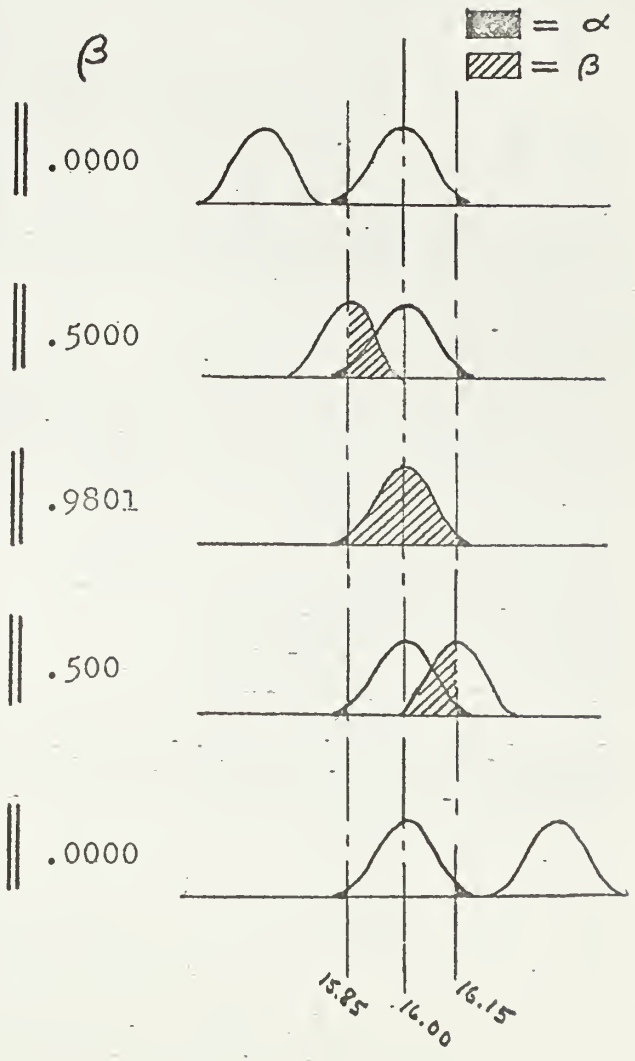
$$u = \frac{15.85 - 15.00}{.4/\sqrt{39}} = \frac{.85}{.0643} = 13.2$$

$$u = \frac{15.85 - 15.85}{.4/\sqrt{39}} = \frac{0}{.0643} = 0$$

$$u = \frac{15.85 - 16.00}{.4/\sqrt{39}} = \frac{-.15}{.0643} = 2.33$$

$$u = \frac{16.15 - 16.15}{.4/\sqrt{39}} = \frac{0}{.0643} = 0$$

$$u = \frac{16.15 - 17.00}{.4/\sqrt{39}} = \frac{.85}{.0643} = 13.2$$



between 15.85 ounces and 16.15 ounces, allow the process to continue. If the $\bar{x} \leq 15.85$ ounces, or $\bar{x} \geq 16.15$ ounces, then stop the machine (accept the alternative) and determine the cause. Figure 17-2 is the power curve and Figure 17-3 is the operating characteristic curve, both derived from data in Table 17-1.

Table 17-1

μ	P(Accepting $H_o \mid \tilde{\mu} = \mu$) $= \beta$	P(Rejecting $H_o \mid \tilde{\mu} = \mu$) $= \alpha$
15.00	.0000	1.0000
15.85	.5000	.5000
16.00	.9801	.0199
16.15	.5000	.5000
17.00	.0000	1.0000

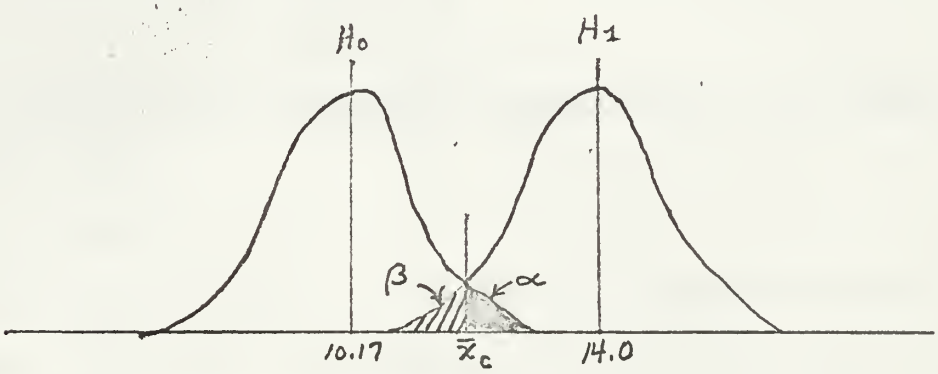
17.7 Controlling Alpha and Beta on the one-tailed test

Suppose a manufacturer of three-inch nylon line for the U. S. Navy has discovered a new manufacturing process and claims that his new process makes better nylon line with a mean breaking strength of 28,000 lbs. normally distributed with a standard deviation of 2,400 lbs. The standard stock three-inch nylon line has a mean breaking

strength of 20,340 lbs. with a standard deviation of 2,400 lbs. The Bureau of Ships decides to test the new line and sets the α and β risks at the .025 level. What should be the size of the sample n and the $\bar{x}_c$ in order to meet the prescribed criteria? Develop the operating characteristic and the power curve for the one-tailed test. Figure 17-4 portrays the problem.

(Note: The breaking strength has been converted to Tons for ease of calculation.)

Figure 17-4



17.8 Computation of Sample size n and $\bar{x}_c$

$$\frac{\bar{x}_c - 10.17}{1.2/\sqrt{n}} = 1.96 \quad \| \alpha = .025 \tag{1}$$

$$\frac{\bar{x}_c - 14}{1.2/\sqrt{n}} = -1.96 \| \beta = .025 \tag{2}$$

Solve (1) and (2) simultaneously for $\bar{x}_c$ and n, from (1)

$$\sqrt{n} = \frac{1.96 \times 1.2}{(\bar{x}_c - 10.17)}$$

and substituting for $\sqrt{n}$ in (2), we have

$$\bar{x}_c - 14 = -(\bar{x}_c - 10.17)$$

$$2.0\bar{x}_c = 24.17$$

$$\bar{x}_c = 12.1$$

then substituting $\bar{x}_c$ in (1), we have

$$\frac{12.1 - 10.17}{1/\sqrt{n}} = 1.96$$

$$\sqrt{n} = \frac{1.96}{1.6}$$

$$n = 1.5$$

but n must be an integer, so

$$n = 2.$$

17.9 Developing the operating Characteristic and Power Curves.

$\bar{x}_c = 12.1$

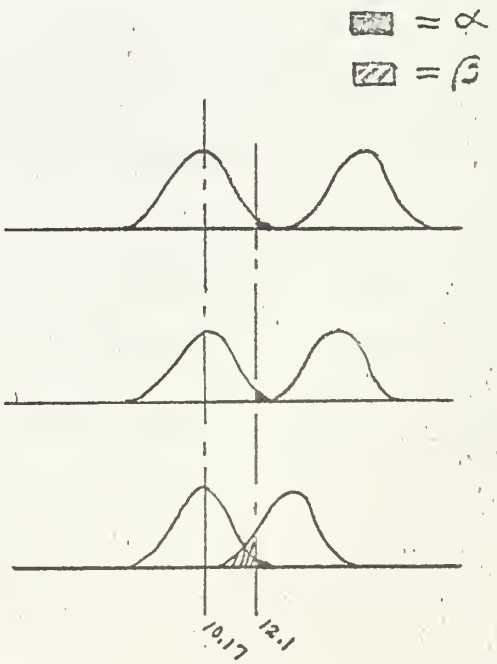
$\sigma_{\bar{x}} = 1.2/\sqrt{n} = .85$

The data in Table 17-2 have been computed using $\bar{x}_c = 12.1$ tons and $\sigma_{\bar{x}} = .85$. Calculations for Table 17-2 and pictorial diagrams are shown below:

$$u = \frac{12.1 - 15}{.85} = -3.41 \quad \| \quad \beta = .0003$$

$$u = \frac{12.1 - 14}{.85} = -2.24 \quad \| \quad \beta = .0126$$

$$u = \frac{12.1 - 13}{.85} = -1.06 \quad \| \quad \beta = .1446$$



$$u = \frac{12.1 - 12}{.85} = .118 \quad \| \quad \beta = .5470$$

$$u = \frac{12.1 - 11}{.85} = 1.295 \| \quad \beta = .9023$$

$$u = \frac{12.1 - 10}{.85} = 2.47 \quad \| \quad \beta = .9968$$

$$u = \frac{12.1 - 9}{.85} = 3.65 \quad \| \quad \beta = .9999$$

$$u = \frac{12.1 - 8}{.85} = 4.92 \quad \| \quad \beta = 1.0000$$

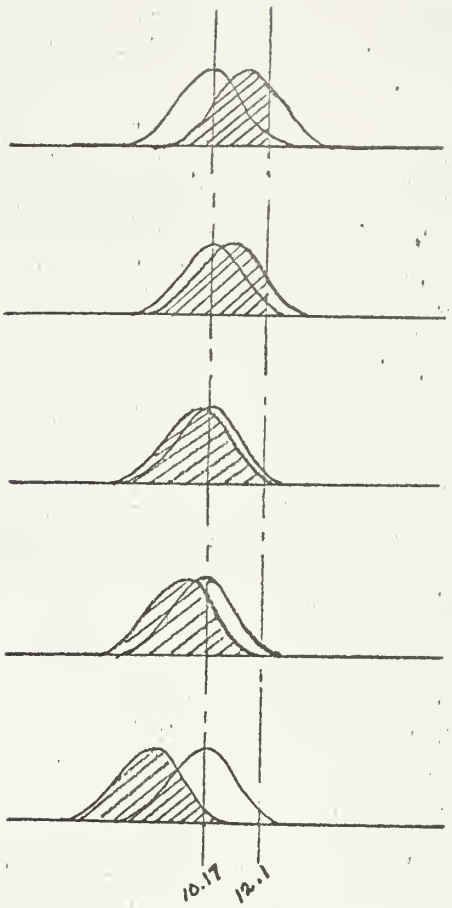
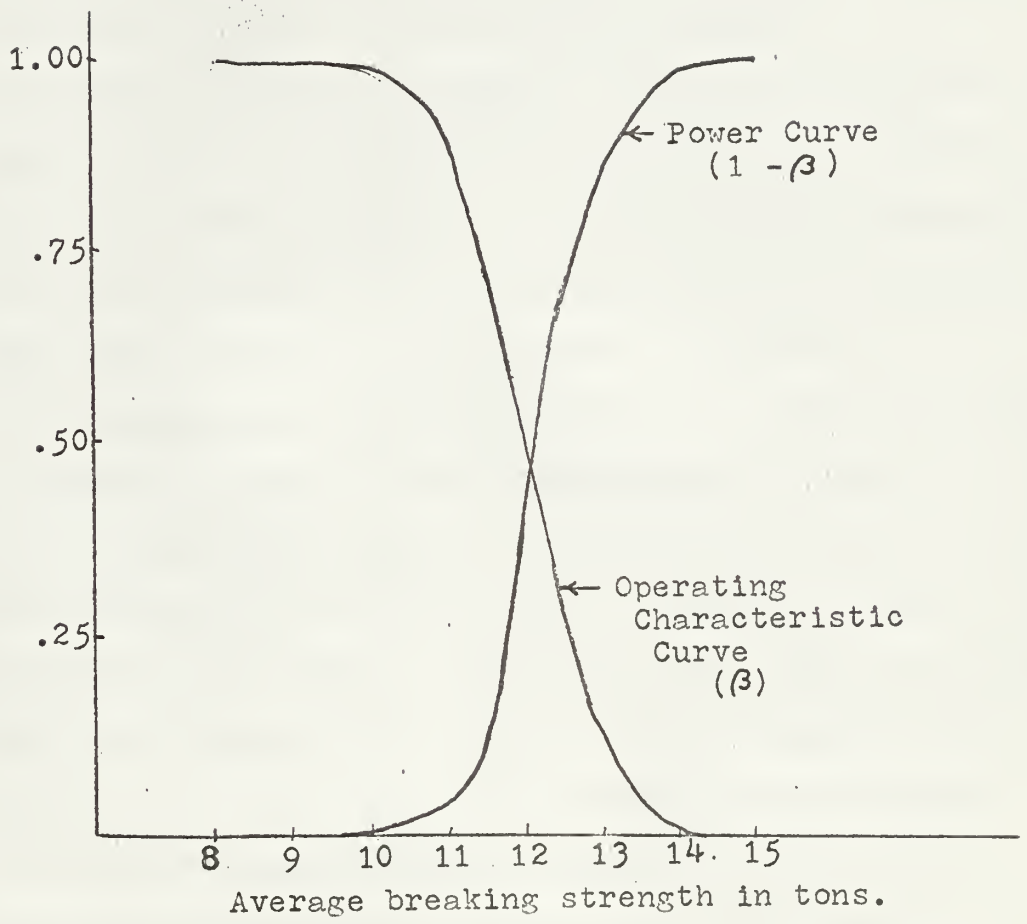


Figure 17-5 is plotted from the above data. From the operating characteristic curve, we see that the probability of keeping the old nylon line when the new breaking strength is less than 10.5 tons is practically 1.0 except for the level of significance of .025 when the new nylon gives a mean of 10.17 tons. It then falls off steeply to zero so that there is practically no chance of continuing to stock the old line when the mean breaking strength exceeds 14 tons, or 28,000 lbs. The power curve is capable of the same interpretation as for the operating characteristic curve, as the two curves are essentially equivalent for this purpose.

Figure 17-5


1.00
.75
.50
.25
Power Curve
(1 - β)
Operating
Characteristic
Curve
(β)
8
9
10
11
12
13
14
15
Average breaking strength in tons.

CHAPTER 18

Expected Loss with Normal Sampling

18.1 General The material presented in Chapter Eighteen of the basic text is essentially the same as that presented in Chapters Ten and Eleven except that we are now using the Normal Distribution instead of the Binomial.

18.2 Three courses of action In making decisions such as whether a batch of material is good and should be used, whether a market exists for a new product, or whether a promotion campaign has been successful, etc., the decision maker can take one of three courses of action:

(1) He can accept without sampling the hypothesis of an affirmative answer to the above questions; that is, that the true μ of the unknown process in which he is interested is beyond the breakeven point. However, by this course he incurs a risk that μ is in fact less than the breakeven point and consequently a loss will be sustained dependent upon how much less than the breakeven point μ actually is.

(2) He can reject without sampling the hypothesis of an affirmative answer to the above questions; that is, that the true μ of the unknown process in which he is interested is below the breakeven point. However, by this course he incurs a risk that μ is in fact greater than the breakeven point and a loss of profit will be sustained dependent upon how much greater than the breakeven point μ actually is.

(3) He can take a sample of the unknown process and thereby attempt to estimate where the true μ actually is--above or below the breakeven point; and by so doing decide whether or not to proceed further with his production, advertising or promotion campaign, etc. However, by taking a sample he not only incurs a risk that the sample is incorrect but he also incurs a cost of sampling which was not incurred by either (1) or (2) above; and the larger the sample, the more expensive its cost.

18.3 The cost of each course of action-Unconditional Expected Total Loss. In deciding which course to take, the decision maker should consider the costs associated with each. These can be computed as follows:

1) Accept without sampling. If the true μ is above the breakeven point, no loss will be sustained, but any μ below the breakeven point will result in a loss whose magnitude is conditional upon the difference between the true μ and the breakeven point. Since the true value of μ is not known and could be any one of a number of values, probabilities are assigned to the occurrence of these various possible values. Then, the conditional loss that would occur for each of these values of μ is multiplied by the probability of its occurrence and results in an expected loss for each respective μ. If we sum these losses we will get the sum of the losses for the

various probabilities of occurrence. This sum, then, is the best estimate of the loss that would be incurred by making an affirmative decision without sampling.

2) Reject without sampling The reasoning here is the same as above except that a negative decision is made without sampling on the assumption that the true μ is below the breakeven point; and therefore if this decision is in error, then a loss of profit will be incurred dependent upon how far above the true μ is from the breakeven point.

3) If a decision is made based on a sample, then the expected loss for each possible value of μ will be the probability that the sample is incorrect multiplied by the conditional loss for that value of μ times the probability, $P(\mu)$, of that μ occuring. The unconditional expected total loss will be the sum of the expected losses plus the cost of sampling.

18.4 The importance of μ From the above it can be seen that the true value of μ determines the risks which a decision maker incurs irrespective of what course of action he chooses to take. This also means that the magnitude of profit or loss is determined in part by the true value of μ.

18.5 The breakeven point In the above discussion considerable mention has been made of the breakeven value in

decision processes. This is the point at which it makes no difference whether the decision maker accepts or rejects the proposition before him or, in other words, the value of μ for which the two possible acts, accept or reject, are equally costly.

18.6 The subjective assignment of $\tilde{\mu}$. Considerable mention has also been made of the probability, $P(\mu)$, of various values of μ occurring. The decision maker, of course, is seeking mathematically objective information on which to base his decision. Yet the values of $P(\tilde{\mu})$ may be quite subjective--they may be based on prior beliefs about a process (or the distribution of μ); or on the other hand, they may be quite accurate, having been based on a collection of historical information. Although generally subjective, the values assigned to $P(\mu)$ should be based on the most accurate information available, and in this connection, later material in the text provides a better method of assigning values to $P(\mu)$.

BIBLIOGRAPHY

Bierman, Harold J., Fornaker, Lawrence E., and Jaedicks, Robert K. Quantitative Analysis for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc., 1961.

Bowen, Earl K. Statistics. Homewood, Illinois: Richard D. Irwin, Inc., 1960.

Freund, John E., Williams, Frank J. Modern Business Statistics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1958.

Griffin, John I. Statistics Methods and Application. New York: Holt, Rinehart and Winston, 1962.

Hirshleifer, Jack. The Bayesian Approach to Statistical Decision, and Exposition.

Kurnow, Ernest, Glasser, Gerald J. and Ottman, Frederick R. Statistics for Business Decisions. Homewood, Illinois: Richard D. Irwin, Inc., 1959.

Miller, David W. and Starr, Martin K. Executive Decisions and Operations Research. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1961.

Mosteller, Frederick, Rourke, Robert I. K., and Thomas, George B., Jr. Probability with Statistical Applications. Reading, Mass: Addison-Wesly Publishing Company, Inc.

Neter, John and Wasserman, William. Fundamental Statistics for Business and Economics. Boston: Allyn and Bacon, Inc., 1961.

Peatman, John G. Introduction to Applied Statistics. New York: Harper & Row, Publishers, 1963.

Perlman, Richard. Problems in Statistics for Economics and Business Students. New York: Holt, Rinehart and Winston, Inc., 1963.

Schlaifer, Robert. Probability and Statistics for Business Decisions. New York: McGraw-Hill Book Company, Inc., 1959.

Smith, Frank, C. and Leabo, D. A. Basic Statistics for Business Economics. Homewood, Illinois: Richard D. Irwin, Inc., 1960.

Spiegel, Murray R. Theory and Problems of Statistics. New York: Schaum Publishing Co., 1961.

Wallis, W. Allen, and Roberts, Harry V. Statistics a New Approach. The Free Press of Glencoe, Inc., 1962.

Wessel, Robert H. and Wilbet, Edward R. Statistics as Applied to Economics and Business. New York: Henry Holt and Company, Inc., 1959.

Yuker, Harold E. A Guide to Statistical Calculations. New York: G. P. Putnam's Sons, 1958.

thesM275
Technical notes on probability and stati
3 2768 002 04177 4
DUDLEY KNOX LIBRARY